DOCUMENT RESUME

ED 394 978 SP 036 694

TITLE	The Implementation of Performance Assessment in Kentucky Classrooms.
INSTITUTION	Louisville Univ., Ky. School of Education.
SPONS AGENCY	Kentucky Inst. for Education Research, Frankfort.
PUB DATE	Aug 95
NOTE	105p.; For related documents, see SP 036 685-694.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Educational Assessment; Elementary Secondary Education; *Performance Based Assessment; Program Evaluation; Program Implementation; Public School Teachers; School Restructuring; State Departments of Education; State Standards; Teacher Attitudes; *Teacher Response; Teaching Experience
IDENTIFIERS	*Kentucky; *Kentucky Education Reform Act 1990

ABSTRACT

This report describes what schools and educators across Kentucky are doing to implement school reform in performance assessment based on the Kentucky Education Reform Act of 1990 (KERA). It provides research-based suggestions about how implementation of programs can be enhanced and how the benefits of reform increased for the Kentucky youth. Thirty-two schools from around the state were randomly selected to participate; they included two elementary schools, one middle school, and one high school from each region of the State. All teachers from the disciplines of language arts, mathematics, science, and social studies completed a performance assessment survey; six teachers from each school were randomly selected for personal interviews using the Performance Assessment Component Configuration (PACC) Map. A total of 500 surveys were analyzed. It was concluded that the extent to which performance assessment is occurring in the classrooms of the teachers selected varied considerably, both within and across schools, in terms of the understanding of what is required of a particular type of assessment and how it should be implemented. KERA support documents are being used by 7 out of 10 teachers surveyed as are multiple forms of assessment, including oral and written open-ended questions, performance events, portfolio assignments, skills tests, and conferencing. In addition, new teachers reported using more performance assessment in instruction than more experienced teachers. Appendices provide: a glossary of terms; the performance assessment component configuration map; performance assessment survey questionnaire; and 14 pages of tables displaying statistical results from the survey. (NAV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 394 978

THE KENTUCKY INSTITUTE FOR EDUCATION RESEARCH

The Implementation of Performance Assessment in Kentucky Classrooms

A Report of Research
conducted by
The School of Education
University of Louisville

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
R. Pankratz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

for the
Kentucky Institute for Education Research
146 Consumer Lane, Frankfort, Kentucky 40601

August 1995

SP036694
BEST COPY AVAILABLE

BOARD OF DIRECTORS

Chair
Ben Richmond
Urban League of Louisville
1535 West Broadway
Louisville, KY 40203-3516

Vice Chair
Gary Dodd
CM Management Services
698 Perimeter Drive, Suite 200
Lexington, KY 40517

Secretary
Robert F. Sexton
The Prichard Committee
for Academic Excellence
P. O. Box 1658
Lexington, KY 40592

Treasurer
Doug Kuelpman
United Parcel Service
1400 North Hurstbourne Parkway
Louisville, KY 40223

Lila Bellando
Churchill Weavers
P. O. Box 30
Berea, KY 40403

Barbara Deeb
WKYU-TV
1 Big Red Way
Bowling Green, KY 42101

Jane Joplin Evans
515 North Main Street
Somerset, KY 42501

Blaine Hudson
439 Strickler Hall
University of Louisville
Louisville, KY 40292

Ernie W. Stamper
Ashland Petroleum Co.
P. O. Box 391
Ashland, KY 41114

Fred D. Williams
70 Pentland Place
Ft. Thomas, KY 41075

Amy Helm Wilson
Murray Ledger & Times
1001 Whitnell Avenue
Murray, KY 42071

Joe Wright
Star Route
Harned, KY 40144

Executive Director
Roger S. Pankratz, Ph.D.
KY Institute for Education Research
146 Consumer Lane
Frankfort, KY 40601

THE KENTUCKY INSTITUTE FOR EDUCATION RESEARCH

146 Consumer Lane • Frankfort, Kentucky 40601 • 502-227-9014 • Fax 502-227-8976

The Implementation of Performance Assessment in Kentucky Classrooms

A Report of Research
Conducted by
The School of Education
University of Louisville

Principal Investigator
Bradley A. Matthews

August 1995

Supported With Funds From
The Kentucky Institute for Education Research

PREFACE

This research project is one of six studies conducted in the spring of 1995 to determine the extent schools and educators across Kentucky had implemented Educational Technology, High School Restructuring, the Primary Program, Professional Development, Performance Assessment and School-Based Decision Making.

The studies were sponsored by the Kentucky Institute for Education Research, supported by funding from The Annie E. Casey Foundation. Each of the research projects was contracted to a Kentucky university that managed the research and employed the services of a team of researchers/field observers, mostly from higher education institutions across the state.

Each study was designed to collect data from a random set of schools across the eight state educational regions. All studies used a research tool developed especially for studying the progress of program implementation called an Innovation Component Configuration Map. The Configuration Map enables researchers to judge the level of implementation of different program components based on a common set of standards and guidelines.

Collectively, through these six studies, more than fifty trained researchers visited 189 schools across the Commonwealth conducting interviews, observing classrooms, training sessions and school council meetings, and reviewing documents and collecting artifacts. To date this research represents the single most comprehensive effort to gage the level of implementation of programs initiated through the Kentucky Education Reform Act of 1990 (KERA).

The Kentucky Institute for Education Research is proud to be able to sponsor these projects and highly commends the members of the research teams and the universities for the excellent work of data collection and analysis they conducted under difficult conditions and a limited budget. On behalf of the Institute, I want to personally express my sincere appreciation to each of the principal investigators for their professional commitment to this statewide effort, their many hours of work beyond those budgeted in the contract and their perseverance to produce a high quality research report.

This report not only describes what schools and educators across the state are doing to implement school reform, it also provides research-based, thoughtful suggestions about how implementation of programs can be enhanced and the benefits of reform increased for the youth of Kentucky.

I sincerely hope you will find the contents of this report both informative and helpful.

Roger Pankratz, Executive Director
Kentucky Institute for Education Research

ACKNOWLEDGMENTS

This study of the implementation of performance assessment in classrooms across Kentucky could not have been completed without the cooperation and contributions of numerous educators in Kentucky. We would like to thank them.

First, we would like to thank the principals and teachers who allowed the field observers to come into their schools and classrooms during a very busy late Winter and early Spring. Snow days and KIRIS testing schedules created a challenge for some of the site visits. The willingness of principals to allow site visits and of teachers to complete survey instruments and participate in interviews is a testimony to the professionalism displayed in these schools.

Second, we would like to thank the field observers who made time in their busy schedules to be trained in the use of the Innovation Configuration Component Map and to travel to schools to survey faculties and conduct focused interviews with randomly selected teachers. The observers are listed on the following page.

We would also like to thank Dr. Gene Hall, University of Northern Colorado, and Carolee Hayes, Douglas County School District in Colorado, for their assistance in the development of the Innovation Configuration Component Map (ICCM) and their helpful suggestions on the general design and implementation of the study. Dr. Archie George, University of Idaho, analyzed the ICCM data and assisted with the interpretation of the Cluster Analysis procedure. I would also like to thank Dr. Carol David, research consultant, for her assistance with analyzing the Performance Assessment Surveys that were completed by the faculties of all participating schools.

I would like to further express my appreciation to Dr. Roger Pankratz, Executive Director of the Kentucky Institute for Education Research, for his tireless effort to pursue relevant evaluation methods such as the ICCM and to help fellow educators with assessing the progress of major strands of the Kentucky Education Reform Act. His guidance, support and encouragement during this study were greatly appreciated.

Finally, I want to express my sincere appreciation to Diane Gossett for her much needed assistance in handling the details of administering the grant and providing helpful assistance on printing and mailing issues.

Brad Matthews, Ed.D.
Director of Curriculum and Assessment
JCPS Gheens Professional Development Academy

FIELD OBSERVERS

Dr. Jim Craig
Western Kentucky University

Dr. Lin Doak
Eastern Kentucky University

Dr. Chelate Grindstaff
Morehead State University

Dr. John Guyton
Murray State University

Dr. Willis Johnson
Murray State University

Dr. Jane Lindle
University of Kentucky

Dr. Ben Malphrus
Morehead State University

Dr. Tony Norman
Western Kentucky University

Dr. Dan Ochs
University of Louisville

Dr. Mary Jo Poole
Alice Lloyd College

Dr. Linda Sheffield
Northern Kentucky University

Dr. Ed Swan
Eastern Kentucky University

Dr. Connie Widmer
Northern Kentucky University

TABLE OF CONTENTS

EXECUTIVE SUMMARY AND RECOMMENDATIONS ix
- The Purpose of the Study ix
- The Statewide Sample ix
- The Data Collection Process ix
- Defining the Innovation: Performance Assessment x
- Conclusions Based on Data Analysis xi
- Recommendations xiii
- Suggestions for Further Research xiv

PURPOSE OF THE STUDY 1
- Background of the Study 1
- Observations from the 1994 KIER Study on the Primary Program 1
- Purposes of the 1995 Performance Assessment Implementation Study 2

DATA COLLECTION PROCEDURES 2
- Study Sample 2
- Observer Training 3
- Protocols 4
- Defining the Innovation: Performance Assessment 4
- Development of the Innovation Component Configuration Map 4
- Development of the Performance Assessment Survey 5
- Artifacts 6
- Analysis of the Data 6

RESULTS OF THE INTERVIEWS AND SURVEYS 7
- A General Pattern of Implementation 7
- Extent and Patterns of Implementation Based on the Configuration Map Data 8
  - A. Frequency of Use of Different Types of Assessment 8
  - B. Content Focus 10
  - C. Quality of Performance Assessments 11
  - D. Relationship of Assessment to Instruction 13
  - E. Teacher Role in Performance Assessment 14
  - F. Student Role in Performance Assessment 15
  - G. Performance Standards 16
  - H. Evaluation and Feedback 17
- Comparison of High and Low Implementors 20
- Teachers with Similar Patterns of Implementation 21
- Results of the Teacher Surveys 25
- Findings Related to School Level 25
- Findings Related to Primary vs. Intermediate Level 26
- Findings Related to Subject Area Taught 26

TABLE OF CONTENTS (cont.)

Findings Related to Years of Teaching Experience 27
Preliminary Analysis of Classroom Assessment Artifacts 27

CONCLUSIONS BASED ON THE ANALYSIS OF THE CONFIGURATION MAP AND TEACHER SURVEY DATA 28

RECOMMENDATIONS 31

SUGGESTIONS FOR FURTHER RESEARCH 33

REFERENCES 34

TABLES

TABLE 1. Percent of Teachers Judged to be Implementing Nine Types of Assessment at Various Levels of Use 9

TABLE 2. Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Content Focus 10

TABLE 3. Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Quality 12

TABLE 4. Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Relationship to Instruction 14

TABLE 5. Percent of Teachers Judged to be Using Different Roles to Implement Performance Assessment 15

TABLE 6. Percent of Teachers Judged to be Utilizing Different Student Roles in Implementing Performance Assessment 15

TABLE 7. Percent of Teachers Judged to be Implementing Performance Standards at Different Levels of Clarity, Congruence, and Relationship to Learning 16

TABLE 8. Percent of Teachers Judged to be Implementing Different Types of Evaluation and Feedback at Different Levels of Use 18

TABLE 9. Spearman Rank-Order Correlations: Overall Average With Items on the Performance Assessment Component Configuration Map 20

TABLE 10. Cluster Analysis of Performance Assessment Component Configuration Map .. 23

APPENDIX A 35
Glossary of Terms 36

APPENDIX B 39
Performance Assessment Component Configuration Map. 40

APPENDIX C 50
Performance Assessment Survey 51

APPENDIX D 53
Table D. 1 Frequency of Use of Open-ended Questions (Oral) by School Level ... 54
Table D. 2 Frequency of Use of Open-ended Questions (Written) by School Level . 54
Table D. 3 Frequency of Use of Performance Event Tasks by School Level. 55
Table D. 4 Frequency of Use of Portfolio Tasks by School Level 55
Table D. 5 Frequency of Use of Culminating Performances by School Level 56
Table D. 6 Frequency of Use of Projects by School Level 56
Table D. 7 Frequency of Use of Conferences by School Level 57
Table D. 8 Frequency of Use of Anecdotal Comments by School Level 57
Table D. 9 Percent of Teachers by School Level Using Selected Support Documents 58
Table D.10 Frequency of Using Technology for Assessment by School Level 59
Table D.11 Frequency of Using Hands-on Techniques for Assessment by School Level 59

TABLES (cont.)

Table D.12 Frequency of Providing Students with Standards in Advance by School Level 60
Table D.13 Frequency of Providing Students with Rubrics by School Level 60

APPENDIX E 61
Table E. 1 Frequency of Use of Open-ended Questions (Written) by Primary and Intermediate Levels. 62
Table E. 2 Frequency of Use of Performance Event Tasks by Primary and Intermediate Levels ..... .62
Table E. 3 Frequency of Use of Projects by Primary and Intermediate Levels 63
Table E. 4 Frequency of Use of Anecdotal Comments by Primary and Intermediate Levels 63
Table E. 5 Frequency of Use of Selected Resource Documents by Primary and Intermediate Levels 64
Table E. 6 Frequency of Providing Students with Standards in Advance by Primary and Intermediate Levels 64
Table E. 7 Frequency of Use of Technology in Assessment by Primary and Intermediate Levels 65
Table E. 8 Frequency of Use of Hands-on Strategies in Assessment by Primary and Intermediate Levels 65
Table E. 9 Frequency of Students Auditing Other Student Work by Primary and Intermediate Levels ..... .66
Table E.10 Frequency of Student Work Displayed by Primary and Intermediate Levels 66

APPENDIX F. 67
Table F. 1 Frequency of Use of Open-ended Questions (Oral) by Subject Area ... 68
Table F. 2 Frequency of Use of Performance Event Tasks by Subject Area 68
Table F. 3 Frequency of Use of Portfolio Tasks by Subject Area 69
Table F. 4 Frequency of Use of Culminating Performances by Subject Area 69
Table F. 5 Frequency of Use of Conferences by Subject Area 70
Table F. 6 Frequency of displaying Student Work by Subject Area ..... .71
Table F. 7 Frequency of Using Technology for Assessment by Subject Area 71

APPENDIX G 72
Table G. 1 Frequency of Use of Performance Event Tasks by Years of Experience . 73
Table G. 2 Frequency of Use of Support Documents by Years of Experience 74

APPENDIX H 75
Table H. 1 Distribution of Collected Performance Assessment Artifacts by Subject Area. 76

THE IMPLEMENTATION OF PERFORMANCE ASSESSMENT IN KENTUCKY CLASSROOMS

EXECUTIVE SUMMARY AND RECOMMENDATIONS

The Purpose of the Study

The Kentucky Education Reform Act (KERA) of 1990 required the Kentucky Department of Education to develop a new statewide assessment system that was "primarily performance based" and to hold schools accountable for student learning. The new school reform law also adopted six Kentucky Learning Goals and required the Council on School Performance Standards to define the six Kentucky Learning Goals in measurable terms. In addition, the State Board for Elementary and Secondary Education was charged with developing a model curriculum framework to guide the design of performance-based learning and to assist schools in using performance assessments in regular instructional programs.

It was the purpose of this study to: (a) determine the extent to which performance assessment was being implemented in the classrooms of selected teachers in 32 randomly selected schools across the state, (b) identify successful implementation patterns and the factors influencing successful practice, (c) develop recommendations for embedding performance assessment in instruction, and (d) refine the Performance Assessment Configuration Component (PACC) Map for performance assessment based on the initial study and suggest ways of using the instrument for self-assessment and as a tool for further research.

The Statewide Sample

Thirty-two schools, four in each of the eight Regional Service Center regions across the state, were randomly selected for this study. A stratified random sampling technique was used to obtain two elementary schools, one middle school, and one high school from each region. Six teachers from the disciplines of Language Arts, Mathematics, Science, and Social Studies were randomly selected at each school for personal interviews using the PACC Map. Additionally, all Language Arts, Math, Science, and Social Studies teachers at each of the schools, including interviewed teachers, were asked to complete a performance assessment survey. One-hundred-ninety-two teachers were interviewed using the PACC Map. While return rates varied from school to school, a total of 500 surveys were analyzed for this study. While the study sample may not be completely representative of the use of performance assessment in classrooms across Kentucky, it is the most complete effort to date to define the components of implementation and to collect data on their use by teachers across Kentucky.

The Data Collection Process

A team of 12 field observers recruited from state colleges and universities was trained to collect information on the implementation of specific components of performance assessment using the PACC Map. The observers conducted focused interviews with each teacher selected

for the study and completed a PACC Map on each. Each interviewed teacher also filled out a survey on performance assessment practices. Principals were requested to distribute the performance assessment surveys to all teachers in their building who taught one of the four core subjects used for this study. Of approximately 600 surveys returned, 500 were considered usable for this study.

The PACC Map contained descriptions of different levels of implementation for 13 components that define the attributes of performance assessment. These components were developed during the Spring of 1994 by a team of school teachers, school administrators, and college professors with an interest and expertise in performance assessment. Early versions of the PACC Map were field tested and subsequent revisions made through the Fall of 1994.

Data collected for analysis included completed PACC Map, notes from the focused interviews, summary observations from the field observers, and artifacts that included sample assessment tasks, rubrics, scoring guides, open-ended questions, and other forms of evaluation. More than 300 artifacts were voluntarily submitted by teachers who were interviewed.

Defining the Innovation: Performance Assessment

Performance assessment, unlike traditional forms of evaluation, focuses not only on what student should know, but also on what they can do with what they know and in more realistic situations or contexts. To be effectively implemented, teachers must examine and, in many instances, change the way they assess student performance in their classrooms. Moving from traditional standardized evaluations to new standard-setting evaluations is a major change in student evaluation. Teachers must not only be knowledgeable about how to use the different types of assessments that are available to them, they must also think about how this form of assessment changes the role that students play in the teaching and learning process.

In assessing the extent of implementation of performance assessment across Kentucky, the research team identified different types of performance assessment, quality issues associated with those performances, and the roles played by the teacher and the student. Eight major components were defined for the study and are listed below:

1. Frequency of Use
2. Content Focus
3. Quality of Performance Assessments
4. Relationship of Assessment to Instruction
5. Teacher Role in Performance Assessment
6. Student Role in Performance Assessment
7. Performance Standards
8. Evaluation and Feedback

Conclusions Based on the Data Analysis

1. The extent to which performance assessment is occurring in the classrooms of the teachers selected for this study varies considerably both within and across schools. Observers found a range of differences in terms of understanding what is required of a particular type of assessment and how it should be implemented. Understanding and utilization of specific innovation components varied from teacher to teacher. There is confusion about the meaning of terms used to describe different types of performance assessment. Terms such as "performance task," "portfolio task," and "culminating performance" are often used interchangeably.

2. KERA support documents such as Transformations, Kentucky Department of Education Content Guidelines, The Kentucky Instructional Results Information System (KIRIS) released, open-response items, and KIRIS Assessment Curriculum Reports are reported being used by at least seven of ten teachers surveyed. How these documents are used varies by individual teacher.

3. KIRIS is having a major impact on the use of performance assessment in the classrooms of teachers at the selected schools. However, the use of performance assessment for many teachers is primarily in preparation for KIRIS tests rather than as an integral part of their daily instruction.

4. Multiple forms of assessment including oral and written open-ended questions, performance events, portfolio assignments, skills tests, and conferencing are used by seven of ten teachers surveyed. More than half used some type of open-ended questions within the scope of an instructional unit.

5. Nine of ten teachers reported using oral and written, open-ended questions on a regular basis and eight of ten teachers reported using portfolio tasks within units of instruction.

6. Seven of ten teachers surveyed reported that they used the KIRIS Assessment Curriculum Report, but a number of interviewed teachers reported that they used only one report at a time and were not able to discuss changes in performance over time. Some teachers confused the KIRIS Assessment Curriculum Report with other KIRIS reports.

7. There is evidence that accountability-grade (non K-3) teachers make more extensive use of performance assessments. Comparing primary teachers (K-3) with intermediate teachers (4-5), intermediate level teachers report:

    - Greater use of open-ended questions and performance events on a daily and weekly basis
    - More often providing students with standards in advance of instruction
    - More frequent use of the instructional strategy where students audit other students' work

8. There were differences in reported use of performance assessments among teachers of Language Arts, Mathematics, Science, and Social Studies:

   - Language Arts and Mathematics teachers use portfolio tasks more extensively
   - Language Arts teachers use student feedback conferences more frequently
   - Mathematics teachers make greater use of national standards
   - Science teachers use performance event tasks and hands-on strategies more frequently
   - Social Studies teachers used textbook materials more for assessment

9. There were differences in reported use of performance assessment among teachers at elementary, middle, and high schools:

   - Primary and intermediate level teachers reported a more frequent use of open-ended oral questions, portfolio assignments, culminating performances, projects, student conferences, anecdotal comments, and use of hands-on assessment activities
   - Primary and intermediate level teachers are more likely to divide instructional time between giving information, coaching students and providing feedback, and tend to provide feedback more on a daily basis than the middle and high school teachers
   - Intermediate level teachers reported a more frequent use of written, open-ended questions and performance events than primary teachers
   - Intermediate level teachers reported a more frequent use of the textbook, KIRIS open-response, released items, and the KIRIS Assessment Curriculum Reports more frequently than primary teachers
   - Middle and high school teachers reported using KIRIS open-response items more frequently than primary or intermediate level teachers
   - High school teachers reported using open-ended written questions, performance events, and technology for assessment activities more than all other levels

10. Differences in the setting of standards for performance assessment were observed among elementary, middle, and high school levels. Elementary teachers display student work most frequently as a standard-setting mechanism. High school teachers are the most likely to provide students with standards in advance and to use scoring rubrics on assessments.

11. High implementors of performance assessment use assessment to drive instruction, use technology and hands-on manipulatives more frequently, provide challenging and engaging assessments, and provide content that covers multiple Kentucky Learning Goals and Academic Expectations. These areas are the greatest predictors for implementation.

12. High implementors of performance assessment use open-ended, written questions and portfolios tasks more frequently. However, types of assessments tend to vary independently of the other components that were measured and by themselves are not good predictors of effective implementation.

13. New teachers report higher uses of performance assessment in instruction than more experienced teachers. For example, 95 percent of teachers with one to five years' experience report using performance events within units of study. The extent of use of performance events drops to 70 percent for teachers in the range of six to ten years and to 50 percent for teachers with more than ten years of teaching experience.

14. Teachers who use KIRIS released items are significantly different from non-users in several ways. They used open-ended written questions and portfolio assignments more frequently than non-users. They also used rubrics for student work and assessment more frequently. Teachers who used KIRIS released items displayed student work more frequently, used the content guidelines more frequently, and used real world examples for student assessment more frequently than non-users.

Recommendations

1. Resource documents should be produced that define the types of assessment and provide rich examples of each type.

2. Professional development should be provided in a variety of areas related to curriculum and assessment. Key areas include:

   a. Design and use of different types of performance assessments that challenge students to perform at higher levels and engage them with meaningful tasks

   b. Design of performance standards and scoring rubrics that elicit quality work

   c. Use of portfolios in more subject areas and as a method for encouraging higher levels of performance for all students through assessment of work in progress

   d. Design and implementation of a standards-based curriculum

   e. Development of quality tasks that can be used to organize and drive instruction

3. Good model lessons, performance assessment tasks, and the methods for developing them should be made available to every teacher.

4. A bank of quality performance assessments should be established by academic expectation and subject area and made available through the Kentucky Educational Television network. Each event or task should also be accompanied by samples of student work that serve as benchmarks for quality.

5. *Transformations: Kentucky's Curriculum Framework* should be updated to incorporate current information in various national standards documents. The Academic Expectations need to be further defined by content standards in order to provide clarity to what students are expected to know and be able to do.

6. Schools need to explore ways of having more of the faculty involved in preparing students at every grade level for the KIRIS assessments.

7. Schools need to develop a multi-year KIRIS Assessment Curriculum Report profile to identify both areas of growth and improvement as well as reporting categories in need of improvement.

8. The Kentucky Department of Education in cooperation with local school districts should identify classrooms and teachers where performance assessment is being used to improve the quality of student work. These teachers and schools should serve as model sites for visits and be involved in the professional development of others.

9. The PACC Map needs to be revised and a training module developed for using the map as a self-assessment tool.

Suggestions for Further Research

Studies need to be conducted to:

1. Replicate the present study findings by having a more extensive implementation study that includes:

    a. A more representative sample of schools and teachers,
    b. Arts and Humanities teachers, and
    c. A comparison of schools meeting or exceeding their thresholds with schools not meeting their thresholds.

2. Identify methods for challenging and engaging students in assigned tasks and student willingness to persist with the assigned task.

3. Identify the extent of implementation within schools and across schools. Factors need to be identified that increase the likelihood of high implementation of performance assessment for all teachers in a school.

4. Identify the effects of professional development associated with the use of performance assessment.

5. Identify the effects of school leadership in promoting the use of performance assessment for instruction.

THE IMPLEMENTATION OF PERFORMANCE ASSESSMENT IN KENTUCKY CLASSROOMS

PURPOSE OF THE STUDY

Background of the Study

In June, 1990, the Kentucky legislature passed the Kentucky Education Reform Act (KERA) mandating a complete overhaul of the Kentucky educational system in the areas of finance, governance, and curriculum. Six Kentucky Learning Goals were adopted as a part of that legislation and the Council on School Performance Standards was charged with defining the six goals in measurable terms. That work resulted in the identification of 68 "Valued Outcomes" across the six Kentucky Learning Goals.

Simultaneously, the Kentucky Department of Education was charged with overseeing the development of a performance-based assessment system that would require students to demonstrate what they know and what they are able to do with the knowledge that they have gained. The beginning phases of the Kentucky Instructional Results Information System (KIRIS) were begun during the 1991-92 school year. In 1994, the legislature revised the scope of KIRIS to include only four of the six Kentucky Learning Goals and required the Kentucky Department of Education to refine the 68 outcomes. Fifty-seven Academic Expectations are the result of that revision.

Observations from the 1994 KIER Study on the Primary Program

During the Spring of 1994, the Kentucky Institute for Education Research (KIER) contracted with the Institute on Education Reform at the University of Kentucky to study the patterns and extent of implementation of the Primary Program that had been mandated by KERA. Elements of performance assessment were studied in sixteen different components of the primary program. While more detailed comparisons will be made later in the paper, overall indications are that primary teachers continue to struggle with the same performance assessment issues identified in the earlier study.

An earlier KIER study on the implementation of the Primary Program in Kentucky (Bridge, 1994) reported that teachers were beginning to use a variety of performance assessments but rarely involved students in self-assessment strategies. Most activities were teacher-initiated and students were actively involved in only about one-half of the classrooms that were observed. Few examples of student work were observed on display.

When compared with middle and high school teachers on these same issues, however, elementary teachers, including primary teachers, generally reported a more frequent effort to implement many practices related to performance assessment. Differences were observed, however, between primary (P1-P4) and intermediate (4-5) teachers. The 1994 research effort studied only primary teachers.

Purposes of the 1995 Performance Assessment Implementation Study

The purpose of this study was to provide an initial picture of the patterns and extent of implementation of performance assessment in classrooms across Kentucky. Specifically, the purposes of the research were:

1. To determine the extent to which performance assessment was being implemented in the classrooms of teachers in thirty-two randomly selected schools geographically distributed across Kentucky

2. To identify successful implementation patterns and the factors influencing successful practice

3. To develop recommendations for further embedding performance assessment in the instructional process

4. To refine the Performance Assessment Configuration Component (PACC) Map for performance assessment based on the initial study and suggest ways of using the instrument for self-assessment and as a tool for further research

DATA COLLECTION PROCEDURES

Study Sample

During March and April, 1995, trained observers visited 32 schools across the eight Regional Service Center Regions of Kentucky. A stratified random sampling technique was used to select two elementary schools, one middle school, and one high school from each region. Because KIRIS testing had focused in the early phase on Language Arts, Mathematics, Science, and Social Studies, six teachers from those disciplines were randomly selected at each middle and high school for personal interviews. Three primary and three intermediate teachers were selected from each of the participating elementary schools for personal interviews. Additionally, all Language Arts, Mathematics, Science, and Social Studies teachers at the selected middle and high schools, including interviewed teachers, were asked to complete a performance assessment survey. All grade level teachers in the selected elementary schools were asked to complete the same survey. The survey contained both demographic questions as well as questions from the PACC Map.

One hundred ninety-two teachers were interviewed and a PACC Map was completed for each teacher. Three hundred and eight additional teachers in the same 32 schools were surveyed using a questionnaire along with the interviewed teachers. A total of 500 surveys were analyzed.

Fifty percent of the 192 teachers interviewed were elementary, 25 percent middle school, and 25 percent high school. Size differences between school levels and the return rates yielded a

and 25 percent high school. Size differences between school levels and the return rates yielded a more even distribution for the questionnaire data. Of the 500 teachers surveyed, approximately 37 percent were elementary teachers, 31 percent middle school teachers, and 32 percent high school teachers.

Comparisons by subject were made for this study. Of the 500 teachers surveyed, approximately 66 percent indicated that they taught a single subject, 28 percent taught more than one subject, and 6 percent gave no response. For those responding teachers where a specific subject area was taught, approximately 21 percent taught Language Arts, 17 percent taught Mathematics, 14 percent taught Science, and 14 percent taught Social Studies.

Teaching experience was represented by a fairly even distribution. Approximately 97 percent of the teachers fell within the 1-30 years of service range. The distribution remained fairly constant, declining slightly in five-year increments until the 25-30 years of service range. The percentage of teachers represented dropped sharply at 26 years of service and beyond.

While return rates varied from school to school, 308 additional, usable surveys were collected from other teachers at these schools, providing a total of 500 surveys for analysis. Data obtained from the survey allowed the researchers to make comparisons between self reports and observer judgements obtained from interviews and allowed a broader comparison across disciplines and levels.

Observer Training

A team of 12 field observers was assembled and trained to collect information on the implementation of specific components of performance assessment using the PACC Map. Nearly all observers were professors from regional universities or colleges in Kentucky. All observers were familiar with performance assessment as an integral part of KERA.

The training of observers occurred during a two-day session in which they reviewed the instrument, refined focused interview questions, and practiced using the PACC Map. Rater differences were discussed in order to gain better insight into what was being described in each component. Finally, research and site visitation protocols were reviewed for the study. Follow-up sessions were conducted with members missing the initial training.

Observers were instructed to make a decision on all components of the map and to record their impressions and anecdotal comments on the map. Each observer was also asked to develop a summary of impressions on the implementation of performance assessment based on the schools they visited and recommendations for improving the PACC Map.

Protocols

Site visit and data collection protocols were developed to ensure uniformity of the sampling process. The superintendent of the district and principal of each randomly selected school were notified by letter from the Kentucky Institute for Education Research requesting permission to schedule site visits. The field observers contacted their selected schools to confirm participation in the study and established the times for visitation and interviews. Six of the original 32 schools declined to participate. Alternate sites were selected from a list randomly generated.

Principals were asked to provide the observer with a listing of faculty members, the subjects they taught, and their planning period. Where possible, the six teachers to be interviewed were selected by the observer from that list. Principals were also asked to provide all appropriate staff members (interviewed and non-interviewed) with a copy of the performance assessment survey prior to the site visit. The observer picked up the surveys at the beginning of each interview and at the end of the site visit. A coding information sheet was attached to each PACC Map and survey to collect demographic information and to ensure confidentiality. Observers also collected voluntarily submitted samples of performance assessments and related artifacts that included sample assessment tasks, rubrics, scoring guides, open-ended questions, and other forms of evaluation. More than 300 artifacts were collected from the interviewed teachers.

Defining the Innovation: Performance Assessment

Performance assessment, unlike traditional forms of evaluation, focuses not only on what student should know, but also on what they can do with what they know and in more realistic situations or contexts. To effectively implement performance assessment, teachers must examine not only the way they assess student performance in their classrooms, but also the nature and quality of the performance expected (Hart, 1994; Schlechty, 1990; Wiggins, 1987, 1989, 1993). Moving from traditional standardized evaluations to new standard-setting evaluations is a major change in student evaluation. Teachers must not only be knowledgeable about how to use the different types of assessments that are available to them, they must also think about how this form of assessment changes the role that students play in the teaching and learning process. Analyzing a performance is more about learning than testing (Wiggins, 1993).

Development of the Performance Assessment Component Configuration Map

The primary research instrument used in this study was developed from a shared conceptual framework for understanding the change process known as the Concerns Based Adoption Model (CBAM). More than twenty years of research on the CBAM model centers around three diagnostic dimensions of the change process: (a) stages of concern, (b) levels of use of the innovation, and (c) innovation configurations. Innovation configurations, as defined by Hall and Hord (1987), focus on the extent to which a new program or practice resembles the intent or ideal of the developer.

The PACC Map was the instrument used in this study to assess the extent to which components of performance assessment have been implemented as proposed by the designers of the Kentucky Education Reform Act (See Appendix A). An initial draft of the instrument was developed in the Spring of 1994 during a week of training and development by a team of school teachers, school administrators, and college professors with an interest and expertise in performance assessment. Early versions of the PACC Map were field-tested, revised, and provided to more than 1,300 elementary, middle, and high schools across Kentucky for review and comment. Feedback from practitioners was used in the Fall of 1994 to revise the implementation maps in preparation for this study.

The PACC Map contains descriptions of different levels of implementation for eight major components. Five of those components have two or more sub-components. The eight major components are:

1. *Frequency of Use*: How often students have an opportunity to use different types of assessment. Nine types of assessment are identified
2. *Content Focus*: Contains two sub-components that describe linkage to standards and breadth of content covered
3. *Quality of Performance Assessments*: Contains three sub-components of authenticity, challenge and developmental appropriateness, and assessment materials
4. *Relationship of Assessment to Instruction*: Defined in terms of the degree to which assessment drives instruction and instruction drives assessment
5. *Teacher Role in Performance Assessment:* Defines facilitation of learning and interaction with students
6. *Student Role in Performance Assessment*: Defines the extent of active engagement students have in the assessment process
7. *Performance Standards*: Contains three sub-components of communication to students, congruence with KIRIS, and the relationship of standards to student evaluation
8. *Evaluation and Feedback*: Defines how often students have an opportunity to experience or use different types of evaluation and feedback

Interview questions were developed for each component and sub-component of the PACC Map to probe the teacher's understanding of the concepts represented and to facilitate the observer's determination of the extent of implementation on a given component.

Development of the Performance Assessment Survey

The Performance Assessment Survey was developed from specific components of the PACC Map (See Appendix B for survey). It was added as a data collection instrument to: (a) gather certain information quickly so that more time could be spent on components requiring follow-up questions, (b) clarify and isolate sub-components not separated on the map, (c) collect additional information not found on the map, and (d) increase the number of teachers analyzed in

the study. Comparison of interviewed and non-interviewed teachers allowed the researchers to evaluate the reliability of the sample obtained for the interview process.

In addition to demographic data, questions related to the use of resource documents particular to Kentucky provided helpful information on the extent to which support material was affecting the implementation of performance assessment in classrooms across the state. Some components such as use of technology and hands-on manipulatives were separated to assess the extent to which each was used in performance assessment. Specific instructional strategies, such as the use of rubrics, were identified and isolated to probe the extent of implementation of a particular component.

Artifacts

Nearly three hundred artifacts related to performance assessment were collected during the site visits. While the extent of voluntary participation varied from school to school and teacher to teacher, an array of samples was collected including 55 performance tasks, 36 rubrics or scoring guides, 31 tests, and 73 open-ended questions. These data were used to estimate the quality of assessments being used as well as to learn more about their purpose, structure, and originality.

Analysis of the Data

Several statistical procedures were employed to analyze the PACC Map data obtained from the interviews. An item analysis was done using the Spearman Rank-Order Correlation to determine which items had high correlations with the overall average. High correlation ratings on specific items helped to determine which items contributed to high implementation of the innovation.

Cluster analyses were performed to identify groups of interviewed teachers who responded similarly on certain items. Eight different clusters were identified. Discriminant analysis procedures were used to identify high and low fidelity users of the innovation as well as factors which distinguished high groups and low groups. Using group membership as the independent variable and PACC Map item rating as the dependent variable, tests of significance were used to identify items which best discriminated between the groups.

Cross-tabulations of item ratings were conducted on the PACC Map and survey data. The additional demographic data and additional questions on the survey provided information useful in explaining observed differences among clusters, groups, and levels.

RESULTS OF THE INTERVIEWS

A General Pattern of Implementation

Based on the results of the configuration map, the interview notes, and the survey, the description of the status of current implementation of performance assessment was developed. It is important to note that this study pulls together the observations of many people and at three different school levels. The teachers who were interviewed and surveyed varied considerably in their understanding and implementation of this innovation. The profiles, therefore, are not indicative of any one classroom or school.

Performance assessment appears to be a complex innovation with several critical components. At this stage of statewide implementation, patterns are more observable at the classroom level rather than schoolwide. Teachers within the same school can vary considerably in their understanding and use of performance assessment. Use of the innovation does not guarantee quality or complete mastery of essential elements. Understanding and using the structures of performance assessment appear to precede the transformation of student learning and evaluation that can result from understanding and embracing a performance-based approach to teaching and learning.

Not surprisingly, KIRIS testing is having an impact on the use of performance assessments in classrooms across the state. More than 60 percent of the teachers interviewed appeared to be using performance assessment on a regular basis. These teachers report using a variety of assessment strategies, but particularly the use of open-ended questions and portfolio tasks or prompts. Performance Event Tasks are often used although less frequently. There is, however, a confusion about the meaning of terms used to describe different types of performance assessment. Terms such as "performance task," "portfolio task," and "culminating performance" were often used interchangeably by interviewed teachers as they described the types of assessment that they were using in their classrooms.

Many teachers still separate the preparation for KIRIS testing from the assessment strategies that are normally used in their classrooms and focus more intently on specific types of tests and strategies for taking the test. Some schools alter the school day to give students more exposure to open-ended questions. Other schools, elementary in particular, have identified an individual teacher who provides students with practice in performance event tasks. Teachers at grade levels where KIRIS testing occurs appear to be engaging in performance assessment more than teachers at grade levels where KIRIS testing does not directly affect them. This seems particularly true at the elementary level. Primary teachers are far less likely to engage in performance assessment activities than their intermediate (grades 4-5) counterparts.

Progress toward the implementation of performance assessment varies considerably on many of the individual innovation components. There is a wide range in the depth of understanding of what certain components require to be effectively implemented. For example, identifying the standards for a particular performance is a critical component of performance

assessment. For this type of assessment approach to be effectively implemented, students should know, in advance, what is expected of them and what good work in this area looks like. While a majority of the teachers interviewed and surveyed indicated that they developed standards for performance, many teachers preferred to explain them orally and did not provide the students with any written description of the standards. When pressed for an explanation, several teachers said that they were not sure if they were supposed to be that specific. One teacher cited KIRIS testing as the reason, saying that the standards were not known in advance or during the testing process.

Although there are few studies with which to compare the classroom use of assessment strategies since the implementation of KERA, performance assessment appears to be a component of KERA that is taken seriously and is being implemented to some degree in most Kentucky classrooms. The differences reported are matters of degree rather than implementation or non-implementation. The extent of implementation and the differences observed are reported in this study by high implementors and low implementors of performance assessment components, grade level, subject area taught, and years of experience. The results of the interviews and survey are organized around the eight major components of the PACC Map.

Extent and Patterns of Implementation Based on the Configuration Map Data

Based on the data obtained and knowledge about what is required for implementation of performance assessment, a dotted line was placed on each component of the configuration map to separate adequate implementation from implementation efforts perceived to be inadequate. Tables 1 to 8 show the percent of teachers judged to be implementing a given component of performance assessment at various levels. Teachers who were rated to the left of the dotted line were perceived to be adequately implementing that component of performance assessment. Teachers who were rated to the right of the dotted line were judged not to be implementing that component of performance assessment in an adequate manner. In many instances, variation A (the first column) has been judged to be the ideal. This is not true in all cases, however, and the reader is cautioned to consider the reality of the statement. For example, in studying the frequency of different types of assessments, the first variation for this study was "daily." It is not reasonable to expect, however, that teachers or students would be involved with "culminating performances" or "projects" on a "daily" basis or that daily would be the most desirable behavior. For this initial study, the ideal behavior has not been identified.

The following descriptions of results for each component of the configuration map are based on teacher's comments during the interviews and observer ratings of perceived performance levels based on those interviews.

Frequency of Use of Different Types of Assessment. Table 1 presents the data on the frequency of use of nine different types of assessment. They are: (a) open-ended questions(oral), (b) open-ended questions (written), (c) performance event tasks, (d) portfolio tasks or prompts, (e) culminating performance, (f) projects, (g) traditional assessment of discrete skills,

TABLE 1

Percent of Teachers Judged to be Implementing Nine Types of Assessment at Various Levels of Use

Open-Ended Questions (Oral)				
A Daily 49%	B Weekly 34%	C End of Unit 5%	D Once or twice a year 5%	E No Response 7%
Open-Ended Questions (Written)				
A Daily 13%	B Weekly 50%	C End of Unit 28%	D Once or Twice a year 5%	E No Response 5%
Performance Event Tasks (On demand, one-hour tasks)				
A Daily 1%	B Weekly 22%	C End of Unit 40%	D Once or Twice a year 22%	E No Response 15%
Portfolio Tasks or Prompts				
A Daily 16%	B Weekly 22%	C End of Unit 40%	D Once or Twice a year 22%	E No Response 15%
Culminating Performance				
A Daily 0%	B Weekly 5%	C End of Unit 62%	D Once or Twice a year 10%	E No Response 23%
Projects				
A Daily 3%	B Weekly 15%	C End of Unit 50%	D Once or Twice a year 20%	E No Response 14%
Traditional Assessment of Discrete Skills				
A Daily 11%	B Weekly 41%	C End of Unit 32%	D Once or Twice a year 7%	E No Response 10%
Conferencing				
A Daily 26%	B Weekly 34%	C End of Unit 16%	D Once or Twice a year 14%	E No Response 9%
Anecdotal Records				
A Daily 9%	B Weekly 22%	C End of Unit 21%	D Once or Twice a year 14%	E No Response 34%

(h) conferencing, and (i) anecdotal comments. The frequency of occurrence was identified as daily, weekly, end of unit, and once or twice a year. The dotted line is a judgement by the researchers regarding an acceptable level of implementation at this stage of performance assessment use across the state.

While the threshold for implementation versus non-implementation varied by the type of assessment, more than half of the teachers interviewed reported using one or more of the assessment types at or above the frequency considered crucial for implementation. During this time period, 84 percent of the interviewed teachers reported using traditional assessment of discrete skills, 76 percent used conferencing strategies, 68 percent used projects, 67 percent used culminating performances, 63 percent used performance event tasks, and 52 percent used anecdotal records.

Content Focus. Table 2 shows the percent of teachers judged to be implementing performance assessment in their classrooms with respect to the content focus of the performance assessment used. Linkage of tasks to Kentucky's Academic Expectations and covering a broad range of content were two sub-components of content focus.

TABLE 2

Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Content Focus

Linkage to Standards			
A	B	C	D
All tasks have a clear link to Academic Expectations	Most tasks have a clear link to Academic Expectations	Some tasks have a clear link to Academic Expectations	Almost no tasks have a clear link to Academic Expectations
18%	47%	25%	10%
Breadth of Content			
A	B	C	D
Assessment tasks address a broad range of content across Academic Expectations in several Learning Goals	Assessment tasks mostly address a range of content across Academic Expectations under one Learning Goal	Assessment tasks address a limited range of content and focus on only a few Academic Expectations	Assessment tasks address a very narrow range of content and focus on one or two Academic Expectations
50%	24%	19%	8%

What content is covered and how content is selected is a major decision for teachers. This component considers the extent to which teachers use the Academic Expectations, Content Guidelines, and national standards documents to shape the scope of the content. While the

Academic Expectations provide a framework for making curriculum decisions, many teachers continue to struggle with what will be taught and what will be left out during the course of a school year. Based on the comments of interviewed teachers, it appears that some teachers are using the Academic Expectations to fundamentally rethink what it is that they will teach and their students will learn. In these instances, the *Transformations* document appears to be a valuable resource for planning and developing the curriculum. Other teachers may refer occasionally to *Transformations* but are more likely to use the state content guidelines as a means of aligning their current curriculum.

Two specific areas were studied under the Content Focus component: (a) Linkage to Standards defined the extent to which teachers were linking their assessments to the Academic Expectations, and (b) Breadth of Content defined the extent to which the content covered and assessed was related to one or more KERA Kentucky Learning Goals and Academic Expectations.

Linkage to Standards. Teachers are generally familiar with the Academic Expectations, especially in their content area. The change from valued outcomes to learner outcomes to Academic Expectations, however, has created some confusion. The extent of linkage of expectations to assessments varies, but 65 percent of the teachers interviewed report that most activities and assessments have a clear link to the Academic Expectations. Some teachers connect every activity in class to an academic expectation. Others use the expectations to organize themes or units of study. Ten percent of the interviewed teachers, however, state that almost none of the assessments that they use have a clear link to the expectations. *Transformations: Kentucky's Curriculum Framework* is the primary source for gaining information about the Academic Expectations and corresponding demonstrators. Some districts and schools have also provided short lists, quick reference guides, or lesson plan outlines that contain these expectations. There are still teachers, however, who report that they do not possess a personal copy of the *Transformations* document or a copy of the Academic Expectations.

Breadth of Content: The nature of the KERA Learning Goals and Academic Expectations makes it possible to develop assessment strategies that cover more than one learning goal or Academic Expectation. For example, a subject area assessed in Goal II might be accomplished in such a way as to require students to use problem-solving strategies (Goal V) or to demonstrate an ability to consider multiple perspectives (Goal VI component). Fifty percent of all interviewed teachers report that they use assessments that address a broad range of content across Academic Expectations and more than one learning goal. An additional 24 percent of the interviewed teachers address a range of content across Academic Expectations under one Learning Goal.

Quality of Performance Assessments. Table 3 presents the percent of teachers using performance assessments with respect to different levels of quality. Quality was defined by: (a) the linkage of the tasks to the real world of the student, (b) the developmental

appropriateness of the task for the learning level of the student, and (c) the use of technology and/or "hands-on" manipulatives where possible.

TABLE 3

Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Quality

Authenticity (content link to the real world of the student)	A	B	C	D
	All assessments present problems and challenges that have a direct link to the real world of the student.	Most assessments present problems and challenges related to the real world of of the student.	Some assessments present real-life applications; however many assessments come from textbook material that presents hypothetical or contrived situations.	Assessments are mostly contrived situations that have little relationship to the student's world of experience
	10%	49%	36%	5%
Challenge and Developmental Appropriateness (challenging, engaging, developmentally appropriate)	A	B	C	D
	Nearly all assessments are challenging and engaging to each student. Assessments stretch the student's performance without requiring more than can be expected.	Most assessments are challenging and engaging. Some assessments appear to be too advanced or too elementary to engage the student.	Some assessments are challenging and engaging; however, most assessments appear like end-of-chapter questions or activities.	Most assessments do not appear to be challenging or engaging to the student. Many seem too advanced or too elementary for the student
	22%	50%	25%	3%
Assessment Materials (use of technology and/or hands-on manipulatives)	A	B	C	D	E
	Most assessments involve the use of technology and/or hands-on manipulatives that engage the student in active learning.	Some assessments involve the use of technology and/or manipulatives that engage the student in active learning.	Occasionally, assessments require the use of technology and/or manipulatives. Most assessments require only paper and pencils to complete tasks.	Assessments generally do not require the use of technology or manipulatives. Textbooks, paper, and pencils are the standard materials needed for assessment.	The student has little or no opportunity to use performance assessment materials.
	21%	33%	26%	19%	2%

Developing quality performance assessments requires not only a thorough understanding of a performance-based approach to evaluation, but also an ability to incorporate essential elements

that motivate the learner. Three specific areas were studied under this component. Authenticity described the extent to which the content taught was linked to the real world of the student. Challenge and Developmental Appropriateness described the extent to which assessment tasks were both challenging and engaging as well as developmentally appropriate. Assessment Materials described the extent of use of various forms of technology and/or manipulatives.

Authenticity. A majority of the teachers interviewed reported that they made an effort to develop authentic assessments. Nearly 60 percent indicated that most of the problems or assessment tasks they developed were related to the real world of the student. A variety of examples were given and many centered around writing or math assignments. Science labs were also given as examples. Understanding of the term "authenticity" appears to vary and simulations are considered by many teachers to be an appropriate example.

Challenge and Developmental Appropriateness. Most teachers reported that they were sensitive to the need for making assessments challenging and engaging as well as developing them at an appropriate level of difficulty. Many also acknowledged, however, that they relied on traditional forms of assessment. Several methods for accomplishing each of these goals were described. Developmental appropriateness was most often achieved through selecting tasks that allowed for varying ability levels or using higher level tasks and then requiring different levels of performance for different students. Interviewed teachers often defined "challenge" in terms of accountability for a grade or use of higher level materials. Student engagement was approached through the use of hands-on activities, various forms of technology, and a general effort to select topics and subjects that teachers perceived were of interest to students.

Assessment Materials. Over half of all interviewed teachers indicated that at least some of their assessment of student performance required some form of technology or hands-on manipulatives. Computer technology was most often used for writing or math portfolio assignments. Access appears to be a problem for some teachers. They report that many computer labs are being used extensively for Language Arts, specifically the writing portfolio. When the labs are available, teachers use them for both assessment and non-assessment activities including typing the results of performance events, drill and practice activities, and games. High school teachers are more likely to use technology for assessment activities than elementary or middle school teachers. The use of hands-on manipulatives for assessment varies considerably. Elementary teachers are more likely to use hands-on materials in assessment than middle school or high school teachers. Science teachers are also more likely to use these materials than teachers of other subject areas.

Relationship of Assessment to Instruction. Table 4 presents the percent of teachers using performance assessment with respect to the degree of relationship between assessment and instruction.

TABLE 4

Percent of Teachers Judged to be Implementing Performance Assessment at Various Levels of Relationship to Instruction

A	B	C	D	E
Teachers always use assessment to drive subsequent instruction, and instruction always drives subsequent assessment.	Teachers largely use assessment to drive subsequent instruction, and instruction largely drives subsequent assessment.	Teachers sometimes use assessment to drive subsequent instruction, and instruction sometimes drives subsequent assessment.	Teachers infrequently use assessment to drive subsequent instruction, and instruction infrequently drives subsequent assessment.	Teachers never use assessment to drive subsequent instruction, and instruction never drives subsequent assessment.
15%	38%	28%	17%	3%

While KERA stresses the interconnection of curriculum, instruction, and assessment, the degree to which that occurs varied widely among interviewed teachers. Over half of the teachers interviewed indicated that they largely used assessment to drive subsequent instruction and instruction to drive subsequent assessment. KIRIS testing, however, has caused many teachers to view assessment as driving instruction. Few examples were given that showed how instruction might drive subsequent assessment. Some teachers indicated that poor student results on a particular assessment might cause them to reteach the content, but they were also likely to note changes for next year as opposed to taking the time to reteach.

The KIRIS Assessment Curriculum Report provided by the Kentucky Department of Education in the Fall of each year offers the most relevant information for assessing student performance on the Academic Expectations. While the document is familiar to seven of ten teachers, fewer teachers indicated that they actually used the data to assess areas of strength and weakness. Seldom were the data on subject area reporting categories compared across two or more years to analyze changes in performance.

Teacher Role in Performance Assessments. Table 5 presents the percent of teachers performing different roles as they interact with students during performance assessment in their classrooms. The various interactive roles include motivating, challenging, facilitating, coaching, giving feedback, and dispensing information. Many teachers are playing multiple roles as they work with students on assessment issues. Facilitator and coach were used to describe teacher roles almost as much as a monitor. Two-thirds of the interviewed teachers indicated that they spent most of their time interacting with students by motivating, challenging, encouraging, inviting reflection, and giving feedback during part or all of the assessment process. On traditional forms of assessment, however, teachers still assumed the monitor function and provided students with very little assistance. Teachers appear to be examining their role in the assessment process and opting for more interaction.

TABLE 5

Percent of Teachers Judged to be Using Different Roles to Implement Performance Assessment

A	B	C	D
The teacher interacts with the student by motivating, challenging, encouraging, inviting reflection and giving feedback.	Most of the time the teacher interacts with the student by motivating, challenging, encouraging, inviting, reflection and giving feedback.	Some of the time the teacher interacts with the student by motivating, challenging, encouraging, inviting, reflection and giving feedback.	Little interaction between the teacher and the student ; teacher is dispenser of knowledge/giver of tests.
23%	43%	25%	9%

Student Role in Performance Assessments. Table 6 represents the percent of teachers establishing different student roles as a means for implementing performance assessment. The roles represent a range of active involvement and responsibility for the student.

TABLE 6

Percent of Teachers Judged to be Utilizing Different Student Roles in Implementing Performance Assessment

A	B	C
The student is a problem-giver, a team-maker, a producer of knowledge, an investigator, and and a user of resources.	The student is supervised by the teacher as a knowledge worker and is viewed as a completer of tasks.	The student is a passive "test-taker" and a completer of assessment items.
32%	54%	13%

A number of teachers reported that students were involved in group work and often assumed roles of team leader, recorder, or project manager. A third of the teachers interviewed indicated that they had their students involved in real problem solving and investigative roles. Over half of the teachers reported that students were involved with completing assessment tasks. A small number of teachers (i.e., 13 percent) viewed students as only passive takers of tests. Based on teacher comments during the interview, students appear to be more actively involved in assessment activities. The work assigned, however, is predominantly teacher generated and most students are more likely to be completers of tasks than problem generators or investigators.

Performance Standards. Table 7 represents the percent of teachers judged to be implementing performance assessment in their classrooms with respect to several issues related to standards. Standards used were defined by three factors: (a) the communication of performance standards to students, (b) the congruence of performance standards with KIRIS models, and (c) the relationship of the performance standards to student evaluation.

TABLE 7

Percent of Teachers Judged to be Implementing Performance Standards at Different Levels of Clarity, Congruence, and Relationship to Learning

Standards of Communication to Students (clarity, examples)			
A	B	C	D
Standards of performance are clearly explained and actively communicated to the student in advance of assessments. A variety of examples showing different levels of performance are on display for the student.	Standards of performance are developed and stated, but, there are few, if any, indicators or examples of student performance giving feedback.	Standards of performance are not fully developed or clearly stated. Performance is judged by the teacher and is mostly subjective.	Standards of performance for student work in school have not been developed.
36%	36%	24%	4%
Standards of Congruence with KIRIS (degree of congruence)			
A	B	C	
Performance standards used in the classroom are completely congruent with standards used in KIRIS scoring guides.	Performance standards used in the classroom have some relationship to the standards used In KIRIS scoring guides.	Performance standards used in the classroom have little or no relationship to standards used in KIRIS scoring guides.	
15%	54%	30%	
Relationship of Performance Standards to Student Evaluation (degree of relationship)			
A	B	C	
The student's work on Performance Assessment is the dominant evaluating factor in reporting student learning progress.	The student's work on Performance Assessment is a contributing factor in evaluating and reporting student progress.	The student's work on Performance Assessment contributes little or nothing to evaluation of the learning progress.	
26%	65%	9%	

Well-developed performance assessments not only specify what is to be accomplished, they also define how well the work must be accomplished. The standards are "clearly articulated and compelling" (Schlechty, 1990), contain examples of quality work, and are known in advance by everyone. Three specific areas were studied under this component. Standards of Communication to Students described the extent to which performance standards were

developed, known in advance, and supported by samples of quality work. Standards of Congruence identified the degree of congruence between the teacher's standards and those used in the KIRIS scoring guides. Relationships of Performance Standards to Student Evaluation described the degree to which students' work on performance assessments determined the report on their progress.

Standards of Communication to Students. Most teachers recognized a need for communicating performance expectations to students. Nearly three-fourths of all the teachers interviewed indicated that they developed and stated performance standards for student work. A third of those same teachers reported that they provided examples of what good student work looked like. Many teachers are learning about or are using rubrics as a way of communicating performance expectations. Some teachers also report a reluctance to provide students with direct examples for fear that they will mimic the work presented and not do their own thinking. Many teacher examples of standards are course or unit expectations of performance rather than standards for specific performance tasks. A good deal of communicating of standards is still done orally as opposed to providing students with hard copy.

Standards of Congruence with KIRIS. Over half of the teachers interviewed indicated that they had referred to the performance standards outlined in the KIRIS scoring guides, but only 15 percent reported using them on a regular basis. When specific use was mentioned, the open response scoring guides and writing portfolio scoring guides were the most frequently mentioned.

Relationship of Performance Standards to Student Evaluation. Nine out of ten teachers interviewed indicated that students' work on performance assessments was at least a contributing factor in evaluating and reporting their progress. A fourth of those same teachers indicated that it was the dominating factor in reporting student learning progress. There appears to be a blending of old and new practices as teachers explore new ways to evaluate student performance. Point systems and grade averages exist alongside of student exhibitions of mastery as teachers try to accommodate new forms of evaluation.

Evaluation and Feedback. Table 8 presents the data on the frequency of use of five types of evaluation and feedback: (a) the extent to which students evaluate or reflect on their own work, (b) the extent to which students audit other students' work. (c) the extent to which teachers evaluate work and give feedback, (d) the extent to which teachers cooperate with the student in the evaluation of student work, and (e) the extent to which teachers display student work.

When performance assessment is properly implemented, evaluation and feedback occur in several ways. The teacher and the student differentiate between work in progress and finished

TABLE 8

Percent of Teachers Judged to be Implementing Different Types of Evaluation and Feedback at Different Levels of Use

	A Daily	B Weekly	C End of Unit	D End of Grading Period	E No Response
Students evaluate and reflect on own work	42%	41%	9%	3%	5%
Students audit other students' work	11%	51%	21%	5%	13%
Teacher evaluates or audits student work and gives feedback	61%	30%	8%	1%	0%
Evaluation is a cooperative effort between teacher and student	28%	35%	17%	7%	12%
Student work is displayed	12%	42%	30%	9%	7%

work and the evaluation or feedback provides the student with both bearing and direction. The type and frequency of five specific types of evaluation and feedback were studied.

Student Self-Evaluation and Reflection. Based on the interviews, most teachers allow or encourage students to evaluate and reflect on their own work. Eighty-three percent of the teachers interviewed indicated that they used this approach on a daily or weekly basis. The type of self evaluation varies and may include grading one's own paper, checking against a rubric, proofreading, asking students to write an evaluation of their own work, or correcting work based on comments from other students or the teacher.

Students' Audit Other Students Work. Many teachers are encouraging or requiring students to audit other students' work. More that 60 percent of the teachers interviewed indicated that they used peer review on a daily or weekly basis. Over 80 percent indicated that they used peer review within a unit of study. Much of this peer auditing occurs around writing assignments. In addition to students checking other students' work, a number of teachers are having students work in pairs to read each other's work or to ask each other focusing questions. This occurs more frequently above the primary level.

Teacher Evaluation and Feedback. Nine out of ten teachers who were interviewed provide students with some form of evaluation and/or feedback on a daily or weekly basis. In addition to routine grading of papers, many teachers are taking the time to write comments to students and conference with students about their work. Interviewed elementary teachers are using conferencing strategies more frequently than either middle school or high school teachers.

Evaluation as a Cooperative Effort. Many of the interviewed teachers report that they conference with the student about the student's work. While most teachers seem to make the final decision about a student's grade, some teachers did report that they might allow the conference to change their position on a particular evaluation. Elementary and middle school teachers used this strategy more frequently than high school teachers.

Display of Student Work. Eight out of ten interviewed teachers reported that they had displayed some student work. The display often tended to be more for total classroom and less as a strategy for providing exemplary models. Some teachers expressed discomfort with publicly identifying and singling out individual work of students they were presently teaching. Others indicated that they would do so anonymously and often only with other classes. Few teachers indicated that they had saved exemplary work from previous years to use as models. Finally, elementary teachers displayed student work more frequently than middle or high school teachers.

Comparison of High and Low Implementors

In an effort to identify which sub-components of performance assessment were most critical to high implementation, Spearman Rank-Order Correlation Coefficients were computed for each of the 25 sub-components on the PACC Map. These are presented in Table 9. Those components with higher correlation coefficients were the most indicative of overall high implementation. In other words, those items with high correlation were the sub-components that separated high implementors from low implementors. Sub-components with low correlation coefficients were poor predictors of high implementation.

TABLE 9

Spearman Rank-Order Correlations: Overall Average With Items on the Performance Assessment Component Configuration Map

Component/Sub-component	Correlation Coefficient	Probability
Relationship of Assessment to Instruction	0.66861	0.0001
Assessment Materials	0.63599	0.0001
Challenge and Developmental Appropriateness	0.62899	0.0001
Breadth of Content	0.62251	0.0001
Teacher Role in Performance Assessment	0.59265	0.0001
Authenticity	0.58608	0.0001
Student Role in Performance Assessment	0.58411	0.0001
Linkage to Standards	0.56552	0.0001
Performance Standards	0.53768	0.0001
Evaluation is a Cooperative Effort	0.53342	0.0001
Standards of Congruence with KIRIS	0.51686	0.0001
Conferencing	0.48625	0.0001
Portfolio tasks/Prompts	0.48377	0.0001
Teacher Evaluates/Audits and Gives Feedback	0.42865	0.0001
Open-ended Questions (Written)	0.40682	0.0001
Relationship of Performance Stds. to Student Evaluation	0.38627	0.0001
Performance Event Tasks (one-hour on-demand tasks)	0.37744	0.0001
Student work is displayed	0.36441	0.0001
Anecdotal Comments	0.36170	0.0001
Open-ended Questions (Oral)	0.35578	0.0001
Students Audit Other Students' Work	0.29540	0.0001
Projects	0.29321	0.0001
Students Evaluate/Reflect on Own Work	0.28592	0.0001
Traditional Assessment of Discrete Skills	0.23306	0.0011
Culminating Events	0.13623	0.0595

n=192

From Table 9, it is evident that *Relationship of Assessment to Instruction, Assessment Materials, Challenge and Developmental Appropriateness*, and *Breadth of Content* are the most critical to high implementation and are the most used by teachers leading the study sample in overall implementation. *Challenge and Developmental Appropriateness* and *Assessment Materials* are two sub-components of *Quality of Performance Assessments*. *Breadth of Content* is a sub-component of *Content Focus*. It should also be noted that *Authenticity*, a third sub-component related to quality, along with *Teacher Role* and *Student Role* in performance assessment were found to be toward the high end of the rank order correlation as well. Types of assessment and evaluation and feedback items showed the least relationship with any other item.

Interestingly, the use of different types of performance assessment is no predictor of high implementation. Low implementors are just as likely to use various performance assessment strategies on a frequent basis as high implementors. This suggests that many teachers are beginning to use a variety of performance assessments but not necessarily attending to the other critical components that can make performance assessment a powerful tool for learning. They are experimenting with the structural aspects of using a performance-based approach for evaluation, but may be struggling with ways to link assessments directly to real-world issues. Preliminary analysis of the artifacts submitted indicates a variety of assessment types are being used; the quality of those assessments also varies considerably.

Teachers With Similar Patterns of Implementation

In an effort to group teachers according to similar patterns of implementation, a Cluster Analysis was used. In this analysis, teachers were grouped into clusters based on two factors: (a) similarity of overall implementation scores, and (b) similarity of patterns of high and low implementation of the various sub-components. The 192 teachers were grouped into Clusters 1 through 7, with Cluster 0 containing a group of outliers that did not fit any of the other clusters.

Table 10 shows the percent of teachers "successfully" implementing each of the 25 sub-components for each of the Clusters 0-7. Some general characteristics of each cluster follow:

Cluster 0- Teacher practices related to performance assessment in this group tend to be scattered across non-critical components and to be low on the components that are represented by high implementation. All three school levels are represented, but the percent of high school teachers exceeds their percent of representation in the study sample.

Cluster 1- Teacher practices related to performance assessment in this group are most characteristic of low implementation. As a group they are least likely to be concerned about issues of challenge and developmental appropriateness, relationship of assessment to instruction, and linkage to standards. They have the fewest percent of teachers successfully implementing 19 of the 25 subcomponents of performance assessment. All three school levels are represented, but the percent of middle school teachers is higher than their percent of representation in the study sample while elementary teachers are under-represented.

Cluster 2- Teacher practices related to performance assessment in this group are mostly the characteristic of high implementation. As a group they are more likely to be concerned about issues of challenge and developmental appropriateness, relationship of assessment to instruction, and linkage to standards. All three school levels are represented, but the percent of elementary school teachers exceeds their percent of representation in the study sample and middle school teachers are under-represented.

Cluster 3- Teacher practices related to performance assessment in this group have varying characteristics of high implementation, but as a group correlate less strongly with items such as portfolio tasks and prompts and evaluation and feedback issues. All three school levels are represented, but the percent of elementary school teachers exceeds their percent of representation in the study sample while high school teachers are under-represented.

Cluster 4- Teacher practices related to performance assessment in this group have varying characteristics of low implementation, and as a group correlate less strongly with items such as portfolio tasks and prompts and evaluation and feedback issues. No middle schools are represented; consequently, elementary and high school teachers exceed their percent of representation in the study sample.

Cluster 5- Teacher practices related to performance assessment in this group have fairly strong characteristics of high implementation, but as a group correlate the least with items such as portfolio tasks and prompts and evaluation and feedback issues. Elementary school teachers are the dominant population of this group with an under-representation of middle school teachers and no high schools teachers.

Cluster 6- Teacher practices related to performance assessment in this group have varying characteristics of low implementation, but as a group correlate more strongly with items such as portfolio tasks and prompts and evaluation and feedback issues. All three school levels are represented, but the percent of middle school teachers exceeds their representation in the study sample with elementary teachers being under-represented in the population.

Cluster 7- Teacher practices related to performance assessment in this group represent the mid-range of implementation, but as a group correlate highly with items such as portfolio tasks and prompts and evaluation and feedback issues. All three school levels are represented, but the percent of high school teachers exceeds their representation in the study sample with elementary teachers being under-represented.

TABLE 10

Cluster Analysis of Performance Assessment Component Configuration Map
Percent of Teachers Successfully Implementing Performance Assessment Component

COMPONENT	CLUSTER 0	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4	CLUSTER 5	CLUSTER 6	CLUSTER 7
OPEN-ENDED QUESTIONS (ORAL)	69	69	97	90	100	100	89	74
OPEN-ENDED QUESTIONS (WRITTEN)	32	34	92	84	71	76	84	42
PERFORMANCE EVENT TASKS	42	48	78	68	29	71	84	72
PORTFOLIO TASKS/PROMPTS	84	79	93	99	100	92	94	79
CULMINATING PERFORMANCE	58	45	74	79	100	62	89	63
PROJECTS	53	48	78	73	85	38	84	84
TRAD. ASSESS Discrete SKILLS	85	83	73	79	100	85	84	94
CONFERENCING	43	76	86	95	100	92	72	69
ANECDOTAL COMMENTS	32	41	67	69	47	46	56	47
LINKAGE TO STANDARDS	26	24	100	74	43	77	89	84
BREADTH OF CONTENT	42	24	100	68	100	100	94	100
AUTHENTICITY (REAL WORLD)	42	14	100	58	57	62	50	94
CHALLENGE/DEVELOP APPROPRIATE	43	31	100	85	71	84	83	90
ASSESSMENT MATERIALS	27	13	93	74	58	46	44	79
RELATIONSHIP ASSESS/INSTRUCT	43	10	88	63	14	62	44	79
TEACHER ROLE IN PERF. ASSESS.	42	17	89	73	29	92	95	90
STUDENT ROLE IN PERF. ASSESS.	63	62	86	92	100	100	100	95

TABLE 10 (continued)

Cluster Analysis of Performance Assessment Component Configuration Map
Percent of Teachers Successfully Implementing Performance Assessment Component

COMPONENT	CLUSTER 0	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4	CLUSTER 5	CLUSTER 6	CLUSTER 7
PERFORMANCE STANDARDS	48	52	100	95	57	54	78	79
STDS OF CONGRUENCE WITH KIRIS	37	38	93	95	43	84	95	68
REL PERF. STDS. TO STUDENT EVAL.	79	69	100	95	100	100	100	100
STUDENTS EVAL/REFLECT OWN WORK	68	65	96	95	100	100	83	73
STUDENTS AUDIT OTHERS WORK	47	79	93	94	100	85	89	85
TCHR EVAL/GIVES FEEDBACK	74	93	100	100	100	100	78	89
EVAL IS COOPERATIVE EFFORT	37	76	94	95	99	92	61	95
STUDENT'S WORK IS DISPLAYED	67	67	93	100	100	92	67	100

A comparison of Cluster 2 (high implementors) with Cluster 1 (low implementors) highlights, again, the areas that distinguish successful implementation of performance assessment. It would appear that higher implementors are teachers who view curriculum, instruction, and assessment as interconnected are thinking through the deeper issues of performance assessment. They are more likely to think about how to assess a range of Learning Goals and Academic Expectations and how to use a variety of materials to engage and challenge students. Use of a type of assessment strategy is not seen as an end in itself, but rather as a means by which students explore the richness of their discipline and make connections between what is taught and the real world. They are experimenting with new roles and relationships between the teacher and the student and expecting students to be more actively involved in the learning process. While teachers in Cluster 2 are outperforming teachers in Cluster 1 in every component but one (traditional assessment of discrete skills), there is less difference between these two groups when comparing the frequency of use of performance assessments and types of evaluation and feedback. These two areas provide the least information in determining successful implementation.

RESULTS OF THE TEACHER SURVEYS

Primary and grade level teachers in elementary schools and Language Arts, Mathematics, Science and Social Studies teachers in middle and high schools selected for this study were asked to complete a 28-item Performance Assessment Survey (see Appendix B). Usable completed surveys were obtained from all 192 teachers interviewed and 308 teachers not interviewed. Tables C.1 to C.14 in Appendix C present the significant findings of the self-reported results of these surveys with respect to: (a) school level, (b) primary and intermediate level, (c) subject area taught, and (d) years of teaching experience. These results are based on completed surveys from 188 elementary teachers, 148 middle school teachers, and 164 high school teachers.

Findings Related to School Level (Appendix D)

1. Elementary teachers report using oral, open-ended questions significantly more frequently than middle or high school teachers.

2. High school teachers report using open-ended, written assignments significantly more often than middle or elementary school teachers.

3. More than half of all teachers report using performance events within units of instruction with high school teachers reporting a higher level of use than elementary or middle school teachers.

4. At least 80 percent of all teachers report using some type of portfolio task within a broad unit of instruction. However, elementary teachers lead their colleagues in middle and high school in the use of portfolio assignments as well as in culminating projects.

5. Four of five elementary teachers report using student conferences as a feedback tool within an instructional unit compared to only two of three middle and high school teachers reporting the use of this strategy.

6. High school teachers report more use of educational technology for assessment than elementary or middle school teachers.

7. Four of five elementary teachers report frequent use of hands-on techniques in the assessment of learning compared to only one of two middle and high school teachers reporting frequent use of this assessment strategy.

8. High school teachers are more likely to provide students with standards and scoring rubrics in advance of assessment tasks than middle school teachers. Middle school teachers are more likely to provide standards and rubrics in advance of assessment than elementary teachers.

9. Nineteen of twenty elementary teachers report displaying of student work sometime during a unit of study. Seven of ten middle school teachers and one of two high school teachers report using this practice.

Findings Related to the non-accountability Grades P1-P4 of the Primary Level vs. the accountability Grade 4 and 5 of the Intermediate Level (Appendix E)

10. Intermediate teachers report using different types of assessment including written, open-ended questions, and performance event more frequently than primary teachers.

11. Intermediate teachers report they provide their students with standards for quality of work expected more frequently than primary teachers.

12. Intermediate teachers have students audit other students' work as a feedback technique for assessment more frequently than primary teachers.

Findings Related to Subject Area Taught (Appendix F)

13. Language Arts teachers report using portfolio assignments, culminating performances, and student conferences to a greater extent than Math, Science, or Social Studies teachers.

14. Science teachers report using performance events to a greater extent than Language Arts, Math, or Social Studies teachers.

15. Language Arts and Mathematics teachers are more likely to use technology for assessment purposes than Science and Social Studies teachers.

16. Science teachers use hands-on techniques for assessment more frequently than Language Arts, Mathematics, and Social Studies teachers.

17. Language Arts teachers display student work more frequently than Math, Science, or Social Studies teachers.

Findings Related to Years of Teaching Experience (Appendix G)

18. Teachers who have entered the workforce since KERA use performance assessments to a greater extent than teachers who were teaching before the law was enacted. Nineteen of twenty teachers with one to five years of experience report using various types of performance assessment. For teachers with six to ten years of experience, seven of ten teachers report use of performance assessments. For teachers with more than ten years of experience, the use of performance assessments in instruction drops to one in two teachers.

19. Recently trained teachers with few years of experience are more likely to refer students to published national standards than more experienced teachers. As years of teaching experience increases, teachers report they are less likely to reference national standards.

Preliminary Analysis of Classroom Assessment Artifacts

During the interview process, teachers were asked if they would be willing to share any assessment related items such as open-ended questions, performance events, tests, rubrics, work guidelines, or grading criteria. Three hundred and twenty-four assessment-related items were collected and analyzed as a result of that request. Due to the voluntary nature of the collection process, the artifacts do not represent a cross section of the interviewed teachers; most items are from the middle or high school level. Some teachers shared more information than others. The observations that follow are intended only as descriptions of what has been collected and do not necessarily represent a complete picture of the quality of assessment that is occurring in classrooms across the state.

In assessing the different types of material collected, it was necessary to combine several category types. Unless the teacher was very specific, the researchers could not always separate open-ended questions from portfolio writing prompts or assessment tasks from portfolio tasks. Culminating performance examples were also combined with project examples. Table G.1 in Appendix G provides a specific breakdown on the frequency of assessment types collected and their representation within the four disciplines of Language Arts, Mathematics, Science, and Social Studies.

The researchers collected 73 open-ended questions, 55 performance tasks, 31 tests (short answer or multiple choice), 15 portfolio prompts, 14 projects/culminating performances, and 136 scoring rubrics or work guidelines. While the number of items varied in each category by subject

area, all subject areas had examples in each area except portfolio tasks/prompts. All 15 examples in this area were in Mathematics because of being specifically labeled as portfolio items.

The quality and complexity of the collected items vary considerably. The assessment-related items seem to reflect different levels of understanding as to what constitutes a particular type of assessment. Both commercially prepared items and tasks were submitted as well as teacher-developed items and tasks. Open-ended questions were presented more frequently than other forms of assessment and their quality tends to be fairly high. Simple word problems and essay questions were also submitted as examples of open-ended questions. Tasks tended to vary in length from one-hour, on-demand tasks to tasks that would require several days to a week. Some tasks were open-ended while others were lab exercises from textbooks or workbooks.

A variety of rubrics and work guidelines were collected. They varied considerably in the way they were structured and the amount of information that they provided the student. The degree of information provided to the student ranged from very specific levels of performance with exacting descriptions of quality to general ranges of behavior. Guidelines for work were also submitted that provided some descriptions of quality and time frames for work completion.

CONCLUSIONS

1. Performance assessment is occurring in the classrooms of the teachers selected for this study. Many teachers are using various types of performance assessment strategies. The extent to which these forms of assessment are occurring, however, varies considerably both within and across schools. Many teachers are working on the structural design issues of performance assessment, but few teachers understand how to use performance assessment to improve the teaching/learning process. There still remains confusion about terms such as "performance task," "portfolio task," and "culminating performance." Teachers often use them interchangeably.

2. KERA support documents such as Transformations, KDE Content Guidelines, KIRIS released, open-response items, and KIRIS Assessment Curriculum Reports are reported being used by at least seven of ten teachers surveyed. How these documents are used varies by individual teacher.

3. KIRIS directly influences the way performance assessment is being used in the classrooms of teachers at the selected schools. Open-ended questions and portfolio prompts and tasks are used in many classrooms. Preparation for the KIRIS test is still seen as a separate issue by many teachers rather than as an integral part of their daily instruction.

4. Multiple forms of assessment including oral and written open-ended questions, performance events, portfolio assignments, skills tests, and conferencing are used by seven of ten teachers surveyed. More than half used some type of assessment within the scope of an instructional unit.

5. Nine of ten teachers report using oral and written open-ended questions on a regular basis and eight of ten teachers report using portfolio tasks within units of instruction.

6. Seven of ten teachers surveyed reported that they used the KIRIS Assessment Curriculum Report. A number of interviewed teachers reported that they used only one report at a time and were not able to discuss changes in performance over time. Some teachers confused the KIRIS Assessment Curriculum Report with other KIRIS reports.

7. There is evidence that accountability grade teachers make more extensive use of performance assessments. Comparing primary teachers (K-3) with intermediate teachers (4-5), intermediate level teachers report:

   - Greater use of open-ended questions and performance events on a daily and weekly basis
   - More often providing students with standards in advance of instruction
   - More frequent use of the instructional strategy where students audit other students' work

8. There were differences in reported use of performance assessments among teachers of Language Arts, Mathematics, Science, and Social Studies.

   - Language Arts and Mathematics teachers use portfolio tasks more extensively
   - Language Arts teachers use student feedback conferences more frequently
   - Mathematics teachers make greater use of national standards
   - Science teachers use performance event tasks and hands-on strategies more frequently
   - Social Studies teachers used textbook materials more for assessment

9. There were differences in reported use of performance assessment among teachers at elementary, middle, and high schools.

   - Primary and intermediate level teachers report a more frequent use of open-ended oral questions, portfolio assignments, culminating performances, projects, student conferences, anecdotal comments, and use of hands-on assessment activities
   - Primary and intermediate level teachers are more likely to divide instructional time between giving information, coaching students and providing feedback, and tend to provide feedback more on a daily basis than the middle and high school teachers
   - Intermediate level teachers report a more frequent use of written, open-ended questions and performance events than primary teachers
   - Intermediate level teachers report a more frequent use of the textbook KIRIS open-response released items and the KIRIS Assessment Curriculum Reports than primary teachers
   - Middle and high school teachers report using KIRIS open-response items more frequently than primary or intermediate level teachers
   - High school teachers report using open-ended written questions, performance events, and technology for assessment activities more than all other levels

10. Differences in the setting of standards for performance assessment were observed among elementary, middle, and high school levels. Elementary teachers display student work most frequently as a standard-setting mechanism. High school teachers are the most likely to provide students with standards in advance and to use scoring rubrics on assessments.

11. High implementors of performance assessment use assessment to drive instruction, use technology and hands-on manipulatives more frequently, provide challenging and engaging assessments, and provide content that cover multiple Kentucky Learning Goals and Academic Expectations. These areas are the greatest predictors for implementation.

12. High implementors of performance assessment use open-ended written questions and portfolio tasks more frequently. However, types of assessments tend to vary independently of the other components that were measured and by themselves are not good predictors of effective implementation.

13. New teachers report higher uses of performance assessment in instruction than more experienced teachers. For example, 95 percent of teachers with one to five years' experience report using performance events within units of study. The extent of use of performance events drops to 70 percent for teachers in the range of six to ten years and to 50 percent for teachers with more than ten years of teaching experience.

14. Teachers who use KIRIS released items are significantly different from nonusers in several ways. They use open-ended written questions and portfolio assignments more frequently than nonusers. They also use rubrics for student work and assessment more

frequently. Teachers who use KIRIS released items displayed student work more frequently, use the content guidelines more frequently, and use real world examples for student assessment more frequently than nonusers.

RECOMMENDATIONS

1. Greater clarity needs to be provided in the use and meaning of different types of performance assessment. Resource documents should be produced or recommended that define the types of assessment and provide rich examples of each type.

2. Professional development is needed in a variety of areas related to curriculum and assessment. Key areas include:

   a. Design and use of different types of performance assessments that challenge students to perform at higher levels and engage them with meaningful tasks

   b. Design of performance standards and scoring rubrics that elicit quality work and serve as guidelines for work in progress

   c. Use of portfolios in more subject areas and as a method for encouraging higher levels of performance for all students through assessment of work in progress

   d. Design and implementation of a standards-based curriculum and strategies for evaluating the current curriculum against those standards

   e. Development of quality tasks that can be used to organize and drive instruction

3. Teachers need good model lessons, performance assessment tasks, and methods for developing them. These resources need to be made available to every teacher. (The Kentucky Department of Education has been reviewing and building a pool of exemplary lessons and units and is in the process of completing a handbook on designing performance tasks.)

4. A bank of quality performance assessments should be established by academic expectation and subject area and made available through the Kentucky Educational Television network. Each event or task should also be accompanied by samples of student work that serve as benchmarks for quality.

5. *Transformations: Kentucky's Curriculum Framework* needs to be updated to incorporate current information in various national standards documents. The Academic Expectations need to be further defined by content standards to provide clarity as to what students are expected to know and be able to do.

6. Schools need to explore ways of having more of the faculty involved in preparing students at every grade level for the KIRIS assessments.

7. Schools need to develop a multi-year KIRIS Assessment Curriculum Report profile to identify both areas of growth and improvement as well as reporting categories in need of improvement.

8. The Kentucky Department of Education in cooperation with local school districts should identify classrooms and teachers where performance assessment is being used to improve the quality of student work. These teachers and schools should serve as model visitation sites and be involved in the professional development of others.

9. The PACC Map needs to be revised and a training module developed for using the map as a self-assessment tool.

SUGGESTIONS FOR FURTHER RESEARCH

1. Replicate the present study findings by having a more extensive implementation study that includes:

   a. A more representative sample of schools and teachers
   b. Arts and Humanities teachers
   c. A comparison of schools meeting or exceeding their thresholds with schools not meeting their thresholds

2. Conduct studies with teachers who are high users of performance assessment to better understand the issues of student challenge and engagement, the ways in which assessment drives instruction and instruction drives assessment and the use of technology and hands-on materials that contribute to student willingness to persist with the task.

3. Conduct studies that explore the extent of implementation within schools and across schools. Factors need to be identified that increase the likelihood of high implementation of performance assessment for all teachers in a school.

4. Conduct studies of the effects of professional development associated with the use of performance assessment. How do structures such as year-long, in-service, collegial support groups, and action research techniques affect the implementation process?

5. Conduct studies of the effects of school leadership in promoting the use of performance assessment for instruction. In what roles can the principal, department chair, team leader and/or instructional coordinator make the greatest contribution toward increasing the use of performance assessment as an integral part of the instruction process?

REFERENCES

Bridge, C. A. (1994). The implementation of Kentucky's primary program. Frankfort, KY: Kentucky Institute for Education Research.

Hall, G. E., & Hord, S. M. (1987). Change in schools: Facilitating the process. New York: State of New York University Press.

Hart, D. (1994). Authentic assessment: A handbook for educators. New York: Addison-Wesley Publishing Company.

Schlechty, P. C. (1990). Schools for the twenty-first century: Leadership imperatives for education reform. San Francisco: Jossey-Bass Publishers.

Wiggins, G. P. (1987). Creating a thought-provoking curriculum. American Educator, 11(4), 10-17.

Wiggins, G. P. (1989). Teaching to the authentic test. Educational Leadership, 46(7), 41-47.

Wiggins, G. P. (1993). Assessing student performance: Exploring the purpose and limits of testing. San Francisco: Jossey-Bass Publishers.

APPENDIX A

Glossary of Terms

Frequency of Use defines how often students have an opportunity to use different types of performance assessment. The nine **Types of Assessment** listed are open-ended (oral), open-ended (written), performance event tasks, portfolio tasks or prompts, culminating performances, projects, traditional assessment of discrete skills, conferencing, and anecdotal records. They are defined as:

Open-ended questions (oral and written) are stimulating prompts or questions that require students to use higher order thinking and allow the use of multiple approaches to solutions.

Performance event tasks are usually short-term tasks (one-hour) that require the student to apply knowledge to a given problem or situation.

Portfolio tasks are tasks or prompts that require a longer period of time to complete and show work in progress documenting learning over time.

Culminating performances are major, end-of-unit or end-of-course performances that require students to demonstrate a thorough mastery of the content presented and an integration of that knowledge across multiple goals.

Projects are usually long-term tasks that serve as the focal point for a unit of study.

Skills tests are short (one hour or less) examinations that test for understanding and application of discrete skills within a discipline.

Conferences are one-on-one conversations between the teacher and student that allow for probing the extent to which a student understands a concept, problem, or issue.

Anecdotal Records are narrative descriptions of a student's behavior or performance in progress.

Content Focus contains two subcomponents that describe linkage to standards and the breadth of the content covered.

Linkage to Standards defines the extent to which assessment tasks are linked to the Academic Expectations.

Breadth of Content defines the extent to which the content of the assessment covers one or more of the six Kentucky Learning Goals established in KERA as well as the number of Academic Expectations.

Quality of Performance Assessments is defined through the three sub-components of Authenticity, Challenge and Developmental Appropriateness, and Assessment Materials.

Authenticity defines the extent to which the content has been linked to real world experiences.

Challenge and Development Appropriateness defines the extent to which the assessments are viewed as challenging and engaging for the student.

Assessment Materials defines the extent of use of technology and/or hands-on manipulatives.

Relationship of Assessment to Instruction is defined in terms of the degree to which assessment drives instruction and instruction drives subsequent assessment.

Teacher Role in Performance Assessments is defined in terms of facilitation and interaction as a range of possibilities from dispenser of knowledge to serving as motivator, encourager and giver of feedback.

Student Role in Performance Assessments focuses on the extent of active engagement from passive test taker to investigator and producer of knowledge.

Performance Standards is defined through the three components of standards of communication to students, congruence with KIRIS, and relationship of performance standards to student evaluation.

Standards of Communication to Students defines the extent to which performance standards are clearly explained in advance of the performance

Standards of Congruence with KIRIS defines the degree of congruence of performance assessments used with the KIRIS scoring guides.

Relationship of Performance Standards to Student Evaluation defines the extent to which student work on performance assessment determines a student's evaluation.

Evaluation and Feedback defines how often students have an opportunity to experience or use different types of evaluation and feedback. The five Types of Evaluation and Feedback listed are: student evaluates and reflects on own work, student audits other student work, teacher evaluation of student work, cooperative effort between student and teacher, and display of student work.

Student Evaluates and Reflects on Own Work defines the frequency with which a student uses self reflection, assessment, editing or rewrites of personal work.

Students Audit Other Students' Work defines the frequency with which students assist other students as reviewers and friendly critics. It is not limited to the grading of another student's paper.

Teacher Evaluates or Audits Students' Work and Gives Feedback defines the frequency with which the teacher reviews work in progress as well as completed, final work assignments. It is not limited to the grading of papers.

Evaluation is a Cooperative Effort Between Students and Teachers defines the frequency with which the teacher allows the student to actively participate in the evaluation process defending positions, offering alternative explanations, demonstrating an understanding of the performance standards associated with the work assigned.

Student's Work is Displayed defines the frequency with which the teacher displays exemplary work that serves as models of what meeting the standards looks like. It is not limited to the routine display of classwork regardless of performance level.

APPENDIX B

Performance Assessment Component Configuration Map

DRAFT 1/3/95

An Innovation Component Configuration Map for Performance Assessment

School:	Observer:	Date:
Total Number of Teachers:	Total Number of Students:	

Please note:

This document was developed by the Kentucky Institute for Education Research for the purpose of studying the implementation of Performance Assessment and is not to be used as an evaluation instrument. While it was designed as a research tool, this document can be used for planning and self–assessment of local patterns of implementation.

This document, known as a Component Configuration Map, identifies key components of Performance Assessment and describes variations in practice one would expect to find across the state. The variations farthest to the left are considered by Kentucky practitioners, researchers, and developers to be the emerging practice advocated in the KERA (Kentucky Education Reform Act) initiative. Determining which is the most effective or efficient variation of practice will be the challenge of ongoing research.

The developers of this Innovation Component Configuration Map are periodically reviewing and revising this instrument to improve its usefulness and its ability to identify important variations in practice. Please send all comments and suggestions to Roger Pankratz, executive director, Kentucky Institute for Education Research, 146 Consumer Lane, Frankfort, Kentucky 40601, or fax to (502) 227–8976.

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

TABLE OF CONTENTS

A. Frequency of Use 2
B. Content Focus 3
1. Linkage to Standards 3
2. Breadth of Content 4
C. Quality of Performance Assessments 5
1. Authenticity 5
2. Challenge and Development Appropriateness 5
3. Assessment Materials 6
D. Relationship of Assessment to Instruction 6
E. Teacher Role in Performance Assessments 7
F. Student Role in Performance Assessments 7
G. Performance Standards 8
1. Standards of Communication to Students 8
2. Standards of Congruence with KIRIS 9
3. Relationship of Performance Standards to Student Evaluation 9
H. Evaluation and Feedback 10

PERFORMANCE ASSESSMENT

Innovation Component Configuration (ICC) Map:

A. Frequency of Use *(How often do students have an opportunity to practice?)*
Note the type(s) of assessment for each time period in both columns and rows.

Type of assessments	(a) Daily	(b) Weekly	(c) End of Unit	(d) Once or Twice a Year
1. Open-Ended Questions (Oral)				
2. Open-Ended Questions (Written)				
3. Performance Event Tasks (On demand, one-hour tasks)				
4. Portfolio Tasks or Prompts				
5. Culminating Performance				
6. Projects				
7. Traditional Assessment of Discrete Skills				
8. Conferencing				
9. Anecdotal Records				
10. Other (specific):				

Interview Question(s) (This information may also be taken from the Survey Instrument: Questions 1–10.)

- How do you determine what you want students to know and do in your class?
- How are you using Performance Assessment in your classroom?
- How do you develop assessment activities for your students?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

BEST COPY AVAILABLE

B. Content Focus *(academic expectations, state content guidelines, related to national standards)*

Circle the statement that most accurately applies.

1. Linkage to Standards

(a)	(b)	(c)	(d)
All tasks have a clear link to Academic Expectations.	Most tasks have a clear link to Academic Expectations.	Some tasks have a clear link to Academic Expectations.	Almost no tasks have a clear link to Academic Expectations.

Interview Question(s) (Some of this information may be taken from the Survey Instrument: Questions 16, 17, and 18.)

- Do Kentucky's Academic Expectations national standards affect how you assess students? If so, how? Can you give an example?

- Do the state or District content guidelines affect how you assess students? If so, how? Can you give an example?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

B. Content Focus *(continued)* ***(Academic Expectations, state content guidelines, related to national standards)***
Circle the statement that most accurately applies.

2. Breadth of Content

(a)	(b)	(c)	(d)
Assessment tasks address a broad range of content across Academic Expectations in several Learning Goals.	Assessment tasks mostly address a range of content across Academic Expectations under one Learning Goal.	Assessment tasks address a limited range of content and focus on only a few Academic Expectations.	Assessment tasks address a very narrow range of content and focus on one or two Academic Expectations.

Interview Question(s) (Part of this information may be taken from the Survey Instrument: Question 19.)

- How do you use *Transformations: Kentucky's Curriculum Framework*? (See also background information on survey.)

- Have you used any of the KIRIS, common-released items from previous years to assist you with developing Performance Assessments? If so, how? (See also background information on Survey.)

- What role does the textbook play in your development of student Performance Assessments?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

BEST COPY AVAILABLE

C. Quality of Performance Assessments

Circle the statement that most accurately applies.

1. Authenticity *(content link to the real world of the student)*

(a)	(b)	(c)	(d)
All assessments present problems and challenges that have a direct link to the real world of the student.	Most assessments present problems and challenges related to the real world of the student.	Some assessments present real-life applications; however, many assessments come from textbook material that presents hypothetical or contrived situations.	Assessments are mostly contrived situations that have little relationship to the student's world of experience.

Interview Question(s)

- How do your assessments model real-world experiences for your students? (See also Survey Questions 20-21.)

2. Challenge and Development Appropriateness *(challenging, engaging, developmentally appropriate)*

(a)	(b)	(c)	(d)
Nearly all assessments are challenging and engaging to each student. Assessments stretch the student's performance without requiring more than can be expected.	Most assessments are challenging and engaging. Some assessments appear to be too advanced or too elementary to engage the student.	Some assessments are challenging and engaging; however, most assessments appear like end-of-chapter questions or activities.	Most assessments do not appear to be challenging or engaging to the student. Many seem too advanced or too elementary to engage the student.

Interview Question(s)

- How do you make Performance Assessments challenging for your students?

- How do you make Performance Assessments engaging for your students?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

C. Quality of Performance Assessments *(continued)*

Circle the statement that most accurately applies.

3. Assessment Materials *(use of technology and/or manipulatives when appropriate)*

(a)	(b)	(c)	(d)	(e)
Most assessments involve the use of technology and/or hands-on manipulatives that engage the student in active learning.	Some assessments involve the use of technology and/or manipulatives that engage the student in active learning.	Occasionally, assessments require the use of technology and/or manipulatives. Most assessments require only paper and pencils to complete tasks.	Assessments generally do not require the use of technology or manipulatives. Textbooks, paper, and pencils are the standard materials needed for assessments.	The student has little or no opportunity to use Performance Assessment materials.

Interview Question(s) (Information may also be taken from the Survey Instrument: Questions 14–15.)

- How do you use technology or hands-on manipulatives in the assessment of student performance?

D. Relationship of Assessment to Instruction *(degree of integration with instruction)*

Circle the statement that most accurately applies.

(a)	(b)	(c)	(d)	(e)
Teachers always use assessment to drive subsequent instruction, and instruction always drives subsequent assessment.	Teachers largely use assessment to drive subsequent instruction, and instruction largely drives subsequent assessment.	Teachers sometimes use assessment to drive subsequent instruction, and instruction sometimes drives subsequent assessment.	Teachers infrequently use assessment to drive subsequent instruction, and instruction infrequently drives subsequent assessment.	Teachers never use assessment to drive subsequent instruction, and instruction never drives subsequent assessment.

Interview Question(s)

- How do you use the KIRIS Assessment Curriculum Report for your school to guide instruction and assessment in your subject area? (See also background information on the Survey and Question 22.)

- How do student results on classroom Performance Assessments influence instructional decisions?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

BEST COPY AVAILABLE

E. Teacher Role in Performance Assessments *(facilitation, interaction)*

Circle the statement that most accurately applies.

(a)	(b)	(c)	(d)
All teachers interact with the student by motivating, challenging, encouraging, inviting reflection, and giving feedback.	Most of the time teachers interact with the student by motivating, challenging, encouraging, inviting reflection, and giving feedback.	Some of the time teachers interact with the student by motivating, challenging, encouraging, inviting reflection, and giving feedback.	Little interaction between the teacher and the student; teacher is dispenser of knowledge/giver of tests.

Interview Question(s) (See also Survey Question 23.)

- As students are engaged in a Performance Assessment, what do you do?
- What kinds of interaction do you have with students *during* the assessment of performance?
- What types of Performance Assessment have been the most useful for you in assessing student performance?
- What type of Performance Assessment is the most difficult for you? Why?

F. Student Role in Performance Assessments *(active engagement)*

Circle the statement that most accurately applies.

(a)	(b)	(c)
The student is a problem-giver, a team-maker, a producer of knowledge, an investigator, and a user of resources.	The student is supervised by the teacher as a knowledge worker and is viewed as a completer of tasks.	The student is a passive "test-taker" and a completer of assessment items.

Interview Question(s)

- What are some roles that your students perform in classroom assessment activities?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

G. Performance Standards

Circle the statement that most accurately applies.

1. Standards of Communication to Students *(clarity, examples)*

(a)	(b)	(c)	(d)
Standards of performance are clearly explained and actively communicated to the student in advance of assessments. A variety of examples showing different levels of performance are on display for the student.	Standards of performance are developed and stated, but there are few, if any, indicators or examples of student performance available.	Standards of performance are not fully developed or clearly stated. Performance is judged by the teacher and is mostly subjective.	Standards of performance for student work in school have not been developed.

Interview Question(s) (See also Survey Questions 11, 12, and 13.)

- How do you communicate the Performance Standards expected on the Performance Assessments that you use?

- Have you ever used former student work as benchmarks for current student work in your classroom? How was that structured?

- Do you have a rubric or set of Performance Standards you have used recently? How did you use them?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

BEST COPY AVAILABLE

G. Performance Standards *(continued)*

Circle the statement that most accurately applies.

2. Standards of Congruence with KIRIS *(degree of congruence)*

(a)
Performance Standards used in assessments in the classroom are completely congruent with standards used in KIRIS scoring guides.

(b)
Performance Standards used in the classroom have some relationship to standards used in KIRIS scoring guides.

(c)
Performance Standards used in the classroom have little or no relationship to standards used in KIRIS scoring guides.

Interview Question(s) (See also Survey Questions 11 and 13.)

- Have you ever used the Performance Standards outlined in the KIRIS scoring guides? If so, how?

3. Relationship of Performance Standards to Student Evaluation *(degree of relationship)*

(a)
The student's work on Performance Assessment is the dominant evaluating factor in reporting student learning progress.

(b)
The student's work on Performance Assessment is a contributing factor in evaluating and reporting student progress.

(c)
The student's work on Performance Assessment contributes little or nothing to evaluation of the learning progress.

Interview Question(s)

- How is student work evaluated in your classroom? How are your grading criteria determined?

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

H. Evaluation and Feedback *(frequency, type of evaluation, and feedback)*

Type of assessments (Note the type(s) of evaluation and feedback observed in each time period in both columns and rows.)	(a) Daily	(b) Weekly	(c) End of Unit	(d) End of Grading Period
1. Students evaluate and reflect on own work.				
2. Students audit other students' work.				
3. Teacher evaluates or audits students' work and gives feedback.				
4. Evaluation is a cooperative effort between teachers and students.				
5. Students' work is displayed.				

Interview Question(s) (Initial information may be taken from the Survey Questions 24-28.)

- What opportunities do you provide for your students to receive evaluation and feedback?

- What types of feedback do you give to your students?

Artifacts
What types of performance tasks have you used this or last year that you would be willing to share with the Kentucky Institute for Educational Research?
(Examples include tasks, rubrics, prompts, scoring guides, planning map, culminating event, etc.)

Developed by the Kentucky Institute for Education Research (KIER).
DO NOT REPRODUCE OR DISTRIBUTE WITHOUT PERMISSION FROM KIER.

APPENDIX C

PERFORMANCE ASSESSMENT SURVEY

The purpose of this questionnaire is to assess the degree of implementation of performance assessment that is occurring in classrooms across the Commonwealth of Kentucky. **No individual teacher comparisons or school comparisons will be made.** Please read each statement carefully. Fill in the blank, check, or circle the answer that most accurately reflects your degree of practice or implementation.

Section I. Background (Check all appropriate items and fill in appropriate blanks)

Primary____ Intermediate (4-5)____ Middle/Junior High ____ High School____

Teaching Experience: 1-5 yrs.____ 6-10 yrs.____ 11-15 yrs.____ 16-20 yrs.____
21-25 yrs.____ 26-30 yrs.____ 31 plus yrs.____

Subject(s) Taught_____

Have you referred to Transformations: Kentucky's Curriculum Framework this year? ___yes ___no
Have you referred to any national standards developed in your subject area this year? ___yes ___no
Have you referred to state or district content guidelines for your subject area this year? ___yes ___no
Have you referred to any common released items from earlier KIRIS assessments this year? ___yes ___no
Have you referred to the KIRIS Assessment Curriculum Report for your school this year? ___yes ___no

---

Section II. Performance Assessment Survey (Check all appropriate items)

How often do you use the following types of performance assessment in your classroom?
(Check appropriate time period for each question)

Type of Assessment	daily	weekly	end of unit	once or twice a year	have not used this year
1. Open ended questions (oral)	___	___	___	___	___
2. Open ended questions (written)	___	___	___	___	___
3. Performance Event Tasks (on demand one hr. tasks)	___	___	___	___	___
4. Portfolio tasks or prompts	___	___	___	___	___
5. Culminating Performance	___	___	___	___	___
6. Projects	___	___	___	___	___
7. Testing of discrete skills	___	___	___	___	___
8. conferencing	___	___	___	___	___
9. Anecdotal Records	___	___	___	___	___
10. Other (specific) _____	___	___	___	___	___

How often have you used the following strategies in your classroom this year?
(Check appropriate time period for each question)

Instructional Strategy	Always	Frequently	Occasionally	Not Yet
11. Students are given the performance standards in advance of a unit, project, or topic.	___	___	___	___
12. Samples of exemplary performance are shown to students or placed on display.	___	___	___	___
13. A rubric is used to evaluate student work.	___	___	___	___
14. Technology is used in assessment process.	___	___	___	___

Instructional Strategy (continued)	Always	Frequently	Occasionally	Not Yet
15. Hands-on manipulative are used in a of a performance assessment.	___	___	___	___

Complete the following sentences by selecting the answer that most accurately reflects your practice.

16. _____ student tasks and assessments in my classroom are linked to Kentucky's academic expectations?
    a. All b. Most c.. Some d. Few e. No

17. _____ student tasks and assessments in my classroom are linked to a set of national content standards?
    a. All b. Most c.. Some d. Few e. No

18. _____ student tasks and assessments in my classroom are linked to the state content guidelines?
    a. All b. Most c.. Some d. Few e. No

19. Student performance assessments in my classroom _____

    a. cover a broad range of content across academic expectations in several learning goals.
    b. address a range of academic expectations under one learning goal.
    c. address a limited range of content and focus on a few academic expectations.
    d. address a narrow range of content and focus on one or two academic expectations.
    e. address content only without regard to any academic expectations.

16. _____student performance assessments in my classroom present problems and challenges that are based in real world experiences.
    a. All b. Most c. Some d. few e. No

17. _____student performance assessments in my classroom come from textbook-related material presented in hypothetical or simulated situations.
    a. All b. Most c. Some d. few e. No

18. _____instructional decisions I make in the classroom are directly related to information gained from the KIRIS Assessment Curriculum Report of my school.

    a. All b. Most c. Some d. few e. No

19. I spend most of the instructional time in my classroom _____

    a. disseminating information, covering the content, imparting knowledge.
    b. mostly disseminating knowledge and content, some coaching, some motivating.
    c. split between imparting knowledge, motivating and challenging students.
    d. mostly coaching, providing feedback, challenging and motivating students.

How often do the following types of evaluation and feedback occur in your classroom?
(Check appropriate time period for each question)

Types of Evaluation/Feedback	daily	weekly	end of unit	end of grading period	not done
21. Students evaluate and reflect on own work	___	___	___	___	___
22. Students audit other students work	___	___	___	___	___
23. Teacher evaluates or audits student work and gives feedback	___	___	___	___	___
24. Evaluation is cooperative effort between student and teacher	___	___	___	___	___
25. Student's work is displayed	___	___	___	___	___

BEST COPY AVAILABLE

APPENDIX D

TABLE D.1

FREQUENCY OF USE OF OPEN-ENDED QUESTIONS (ORAL) PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=183	61.7	30.1	5.5	1.1	1.6
Middle School n=147	33.3	44.2	6.8	4.1	11.6
High School n=163	42.9	37.4	7.4	4.9	7.4

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.77100	1	.00001

TABLE D.2

FREQUENCY OF USE OF OPEN-ENDED QUESTIONS (WRITTEN) PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=181	16.0	45.3	21.5	7.2	9.9
Middle School n=148	3.4	54.7	37.8	2.7	1.4
High School n=164	6.7	62.2	25.6	4.9	0.0

Chi-Square	Value	DF	Significance
Mantel-Haenszel	3.58447	1	.05832

TABLE D.3

FREQUENCY OF USE OF PERFORMANCE EVENT TASKS PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=184	3.8	18.5	40.2	16.3	21.2
Middle School n=146	0.7	15.1	38.4	23.3	22.6
High School n=163	6.7	30.7	27.6	27.6	7.4

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.68725	1	.03039

TABLE D.4

FREQUENCY OF USE OF PORTFOLIO TASKS PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=184	13.6	47.8	26.6	7.6	4.3
Middle School n=148	7.6	35.1	37.8	14.9	4.7
High School n=163	3.1	40.5	42.3	11.7	2.5

Chi-Square	Value	DF	Significance
Mantel-Haenszel	10.63992	1	.00111

TABLE D.5

FREQUENCY OF USE OF CULMINATING PERFORMANCES PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=178	3.4	12.4	65.7	7.9	10.7
Middle School n=145	0.7	6.2	50.3	20.7	22.1
High School n=159	1.3	11.3	39.0	17.6	30.8

Chi-Square	Value	DF	Significance
Mantel-Haenszel	29.83968	1	.00000

TABLE D.6

FREQUENCY OF USE OF PROJECTS PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=181	1.7	18.2	61.9	14.4	3.9
Middle School n=145	3.4	13.1	37.2	34.5	11.7
High School n=160	3.8	11.9	36.9	33.1	14.4

Chi-Square	Value	DF	Significance
Mantel-Haenszel	20.31025	1	.00001

TABLE D.7

FREQUENCY OF USE OF CONFERENCES PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=181	24.3	44.8	13.3	13.3	4.4
Middle School n=144	15.3	31.9	18.8	18.8	15.3
High School n=161	16.8	32.9	15.5	19.9	14.9

Chi-Square	Value	DF	Significance
Mantel-Haenszel	19.31319	1	.00001

TABLE D.8

FREQUENCY OF USE OF ANECDOTAL COMMENTS PERCENT OF TEACHERS BY SCHOOL LEVEL					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Elem. School n=181	23.2	29.8	18.8	8.8	19.3
Middle School n=141	8.5	15.6	9.9	15.6	50.4
High School n=156	5.1	14.1	9.6	11.5	59.6

Chi-Square	Value	DF	Significance
Mantel-Haenszel	76.03184	1	.00000

TABLE D.9

PERCENT OF TEACHERS BY SCHOOL LEVEL USING SELECTED SUPPORT DOCUMENTS					
	Transformations KY Framework	National Standards	Content Guidelines	KIRIS Open-Response Items	KIRIS Assess. Curriculum Rpt
Elem. School n=180	87.2	55.6	83.9	76.1*	68.9
Middle School n=147	90.5	56.5	87.8	89.1*	76.2
High School n=163	90.8	60.7	85.3	91.4*	71.6

Chi-Square	Value	DF	Significance
* Likelihood Ratio	17.93399	3	.00013

TABLE D.10

FREQUENCY OF USING TECHNOLOGY FOR ASSESSMENT PERCENT OF TEACHERS BY SCHOOL LEVEL				
	Always	Frequently	Occasionally	Not Yet
Elem. School n=184	9.8	31.0	27.2	32.1
Middle School n=146	4.8	27.4	46.6	21.2
High School n=161	10.6	40.4	37.9	11.2

Chi-Square	Value	DF	Significance
Mantel-Haenszel	7.67991	1	.00558

TABLE D.11

FREQUENCY OF USING HANDS-ON TECHNIQUES FOR ASSESSMENT PERCENT OF TEACHERS BY SCHOOL LEVEL				
	Always	Frequently	Occasionally	Not Yet
Elem. School n=183	20.2	57.9	20.2	1.6
Middle School n=140	9.3	42.9	32.9	15.0
High School n=162	9.3	37.0	42.0	11.7

Chi-Square	Value	DF	Significance
Mantel-Haenszel	41.94918	1	.00000

TABLE D.12

FREQUENCY OF PROVIDING STUDENTS WITH STANDARDS IN ADVANCE PERCENT OF TEACHERS BY SCHOOL LEVEL				
	Always	Frequently	Occasionally	Not Yet
Elem. School n=182	18.7	34.1	30.2	17.0
Middle School n=144	24.3	44.4	25.0	6.3
High School n=163	39.3	38.7	19.0	3.1

Chi-Square	Value	DF	Significance
Mantel-Haenszel	35.02914	1	.00000

TABLE D.13

FREQUENCY OF PROVIDING STUDENTS WITH RUBRICS PERCENT OF TEACHERS BY SCHOOL LEVEL				
	Always	Frequently	Occasionally	Not Yet
Elem. School n=182	9.9	32.4	35.7	22.0
Middle School n=144	17.6	42.6	31.1	8.8
High School n=163	34.4	36.8	26.4	2.5

Chi-Square	Value	DF	Significance
Mantel-Haenszel	51.95993	1	.00000

APPENDIX E

TABLE E.1

FREQUENCY OF USE OF OPEN-ENDED (WRITTEN) TASKS PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	Once or Twice aYr.	Not Done
Primary n=114	13.2	36.8	23.7	11.4	14.9
Intermediate n=67	20.9	59.7	17.9	0.0	1.5

Chi-Square	Value	DF	Significance
Mantel-Haenszel	18.77474	1	.00001
Contingency Coefficient	.33172		.00017

TABLE E.2

FREQUENCY OF USE OF PERFORMANCE EVENT TASKS PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	Once or Twice aYr.	Not Done
Primary n=117	9.4	38.5	36.8	11.1	4.3
Intermediate n=67	20.9	64.2	9.0	1.5	4.5

Chi-Square	Value	DF	Significance
Mantel-Haenszel	18.77474	1	.00001
Contingency Coefficient	.33172		.00017

TABLE E.3

FREQUENCY OF USE OF PROJECTS PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	Once or Twice aYr.	Not Done
Primary n=116	1.7	25.0	58.6	12.1	2.6
Intermediate n=65	1.5	6.2	67.7	18.5	6.2

Chi-Square	Value	DF	Significance
Mantel-Haenszel	8.12183	1	.00437
Contingency Coefficient	.33172		.02406

TABLE E.4

FREQUENCY OF USE OF ANECDOTAL COMMENTS PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	Once or Twice aYr.	Not Done
Primary n=116	25.0	34.5	21.6	6.0	12.9
Intermediate n=65	20.0	21.5	13.8	13.8	30.8

Chi-Square	Value	DF	Significance
Mantel-Haenszel	9.11476	1	.00254
Contingency Coefficient	.26648		.00784

TABLE E.5

FREQUENCY OF USE OF SELECTED RESOURCE DOCUMENTS PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Transformations KY Framework	National Standards	Content Guidelines	KIRIS Open-Response Items *	KIRIS Assess. Curriculum Rpt *
Primary n=113	87.6	56.6	85.8	69.0	64.6
Intermediate n=67	86.6	53.7	80.6	88.1	76.1

Chi-Square	Value	DF	Significance
* Likelihood Ratio	9.04940	1	.00263
* Phi	.21577		.00379

TABLE E.6

FREQUENCY OF PROVIDING STUDENTS WITH STANDARDS IN ADVANCE PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS				
	Always	Frequently	Occasionally	Not Yet
Primary n=116	14.7	27.6	36.2	21.6
Intermediate n=66	25.8	45.5	19.7	9.1

Chi-Square	Value	DF	Significance
Mantel-Haenszel	11.99730	1	.00053
Contingency Coefficient	.27032		.01162

TABLE E.7

FREQUENCY OF THE USE OF TECHNOLOGY IN ASSESSMENT PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS				
	Always	Frequently	Occasionally	Not Yet
Primary n=117	12.8	28.2	24.8	34.2
Intermediate n=67	4.5	35.8	31.3	28.4

Chi-Square	Value	DF	Significance
Mantel-Haenszel	0.04504	1	.83193
Contingency Coefficient	.16193		.17516

TABLE E.8

FREQUENCY OF THE USE OF HANDS-ON STRATEGIES IN ASSESSMENT PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS				
	Always	Frequently	Occasionally	Not Yet
Primary n=116	25.9	62.1	11.2	0.9
Intermediate n=67	10.4	50.7	35.8	3.0

Chi-Square	Value	DF	Significance
Mantel-Haenszel	17.65847	1	.00003
Contingency Coefficient	.31264		.00018

TABLE E.9

FREQUENCY OF STUDENTS AUDITING OTHER STUDENT WORK PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	End of Grading Per.	Not Done
Primary n=112	12.5	28.6	17.9	5.4	35.7
Intermediate n=64	17.2	51.6	20.3	0.0	10.9

Chi-Square	Value	DF	Significance
Mantel-Haenszel	15.17664	1	.00010
Contingency Coefficient	.31495		.00066

TABLE E.10

FREQUENCY OF STUDENT WORK DISPLAYED PERCENT OF TEACHERS BY PRIMARY AND INTERMEDIATE LEVELS					
	Daily	Weekly	End of Unit	End of Grading Per.	Not Done
Primary n=115	27.8	51.3	18.3	1.7	0.9
Intermediate n=65	15.4	35.4	40.0	4.6	4.6

Chi-Square	Value	DF	Significance
Mantel-Haenszel	13.80916	1	.00020
Contingency Coefficient	.28929		.00248

APPENDIX F

TABLE F.1

FREQUENCY OF USE OF OPEN-ENDED QUESTIONS (ORAL) PERCENT OF TEACHERS BY SUBJECT AREA					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Language Arts n=107	50.5	38.3	2.8	3.7	4.7
Mathematics n=85	41.2	32.9	11.8	3.5	10.6
Science n=68	36.8	47.1	5.9	5.9	4.4
Social Studies n=66	34.8	43.9	4.5	3.0	13.6

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.40007	1	.03594

TABLE F.2

FREQUENCY OF USE OF PERFORMANCE EVENT TASKS PERCENT OF TEACHERS BY SUBJECT AREA					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Language Arts n=107	3.7	15.9	30.8	26.2	23.4
Mathematics n=86	1.2	25.6	30.2	29.1	14.0
Science n=68	8.8	32.4	38.2	13.2	7.4
Social Studies n=66	4.5	27.3	28.8	33.3	6.1

Chi-Square	Value	DF	Significance
Mantel-Haenszel	11.63413	1	.00065

TABLE F.3

FREQUENCY OF USE OF PORTFOLIO TASKS PERCENT OF TEACHERS BY SUBJECT AREA					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Language Arts n=107	12.1	55.1	22.4	8.4	1.9
Mathematics n=86	5.8	30.2	54.7	5.8	3.5
Science n=67	0.0	29.9	34.3	28.4	7.5
Social Studies n=66	1.5	32.4	48.5	13.2	4.4

Chi-Square	Value	DF	Significance
Mantel-Haenszel	24.35730	1	.00000

TABLE F.4

FREQUENCY OF USE OF CULMINATING PERFORMANCES PERCENT OF TEACHERS BY SUBJECT AREA					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Language Arts n=106	0.9	10.4	58.5	15.1	15.1
Mathematics n=83	2.4	6.0	39.8	18.1	33.7
Science n=67	0.0	9.0	44.8	14.9	31.3
Social Studies n=63	0.0	7.9	46.0	19.0	27.0

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.62910	1	.03143

TABLE F.5

FREQUENCY OF USE OF CONFERENCES PERCENT OF TEACHERS BY SUBJECT AREA					
	Daily	Weekly	End of Unit	Once/Twice	Not Yet
Language Arts n=106	18.9	47.2	14.2	12.3	7.5
Mathematics n=83	17.4	20.9	24.4	25.6	11.6
Science n=67	13.8	32.3	9.2	20.0	24.6
Social Studies n=63	10.6	37.9	15.2	24.2	12.1

Chi-Square	Value	DF	Significance
Mantel-Haenszel	7.64552	1	.00569

TABLE F.6

FREQUENCY OF DISPLAYING STUDENT WORK PERCENT OF TEACHERS BY SUBJECT AREA				
	Always	Frequently	Occasionally	Not Yet
Language Arts n=107	24.3	44.9	27.1	3.7
Mathematics n=85	12.9	28.2	44.7	14.1
Science n=67	7.5	49.3	31.3	11.9
Social Studies n=68	10.3	44.1	44.1	1.5

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.53660	1	.03318

TABLE F.7

FREQUENCY OF USING TECHNOLOGY FOR ASSESSMENT PERCENT OF TEACHERS BY SUBJECT AREA				
	Always	Frequently	Occasionally	Not Yet
Language Arts n=105	9.5	41.0	31.4	18.1
Mathematics n=84	8.3	39.3	35.7	16.7
Science n=68	11.8	26.5	47.1	14.7
Social Studies n=67	3.0	23.9	55.2	17.9

Chi-Square	Value	DF	Significance
Mantel-Haenszel	4.45212	1	.03486

APPENDIX G

TABLE G.1

FREQUENCY OF USE OF PERFORMANCE EVENT TASKS PERCENT OF TEACHERS BY TEACHING EXPERIENCE					
	Daily	Weekly	End of Unit	End of Grading Per.	Not Done
1-5 Years n=111	23.3	45.6	26.1	2.8	2.2
6-10 Years n=105	7.9	31.4	32.9	9.3	18.6
11-15 Years n=76	5.2	16.8	29.0	12.9	36.1
16-20 Years n=71	0.0	12.7	39.4	26.8	21.1
21-25Years n=93	1.1	25.8	24.7	25.8	22.6
26-30 Years n=30	0.0	14.4	28.6	14.3	0.0

Chi-Square	Value	DF	Significance
Mantel-Haenszel	11.83280	1	.00058

TABLE G.2

FREQUENCY OF USE OF SUPPORT DOCUMENTS PERCENT OF TEACHERS BY TEACHING EXPERIENCE					
	Transformations KY Framework	National Standards	Content Guidelines	KIRIS Open-Response Items	KIRIS Assess. Curriculum Rpt
1-5 Years n=111	81.1	64.0*	88.3	84.7	67.6
6-10 Years n=105	91.4	58.1*	76.2	82.9	60.0
11-15 Years n=76	90.8	53.9*	88.2	85.5	81.6
16-20 Years n=71	95.8	57.7*	78.9	85.9	81.7
21-25Years n=93	88.2	57.7*	88.2	86.0	75.3
26-30 Years n=30	90.0	43.3*	96.7	76.7	66.7

Chi-Square	Value	DF	Significance
* Mantel-Haenszel	6.93735	1	.00844

APPENDIX H

TABLE H-1

Distribution of Collected Performance Assessment Artifacts by Subject Area

Artifact Type	Writing/ Language Arts	Mathematics	Science	Social Studies
Open-Ended Questions	12	15	27	19
Performance Tasks	5	14	32	4
Traditional Tests	8	5	13	5
Portfolio Prompts/Tasks	0	15	0	0
Culminating Performance/Projects	2	0	5	7
Scoring Rubrics/ Work Guides	31	16	15	19

END

U.S. Dept. of Education

Office of Educational
Research and Improvement (OERI)

ERIC

Date Filmed
October 7, 1996